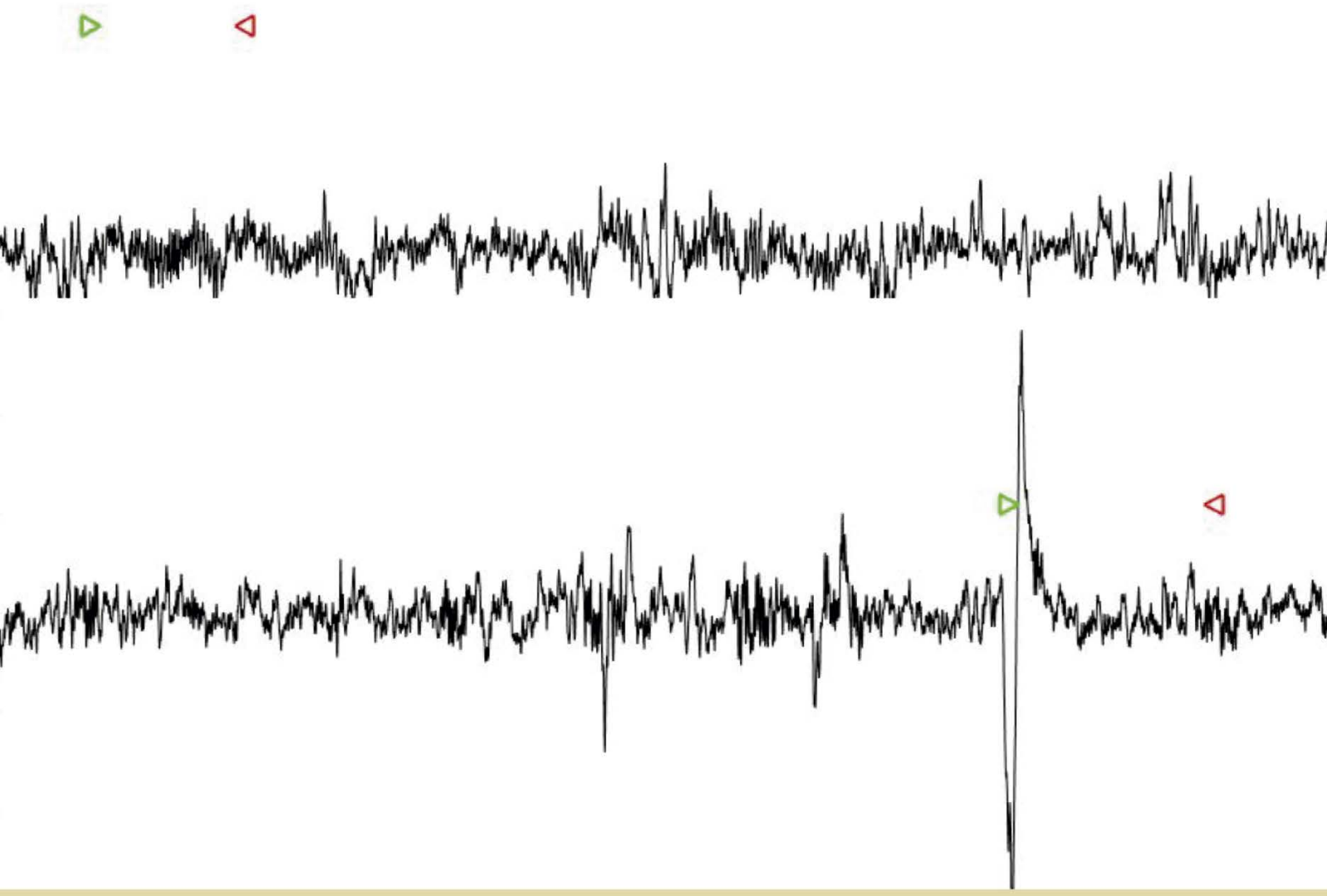


Mariana Duarte Nobre Cortesão

A robust classifier for the automatic detection of transient events in sleep EEGs

Dissertion presented to the University of Coimbra in order to complete the necessary requirements to obtain the Master's degree in Biomedical Engineering.

September 2017

· U C ·

UNIVERSIDADE DE COIMBRA

• U  C •

FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Mariana Duarte Nobre Cortesão

A robust classifier for the automatic detection of transient events in sleep EEGs

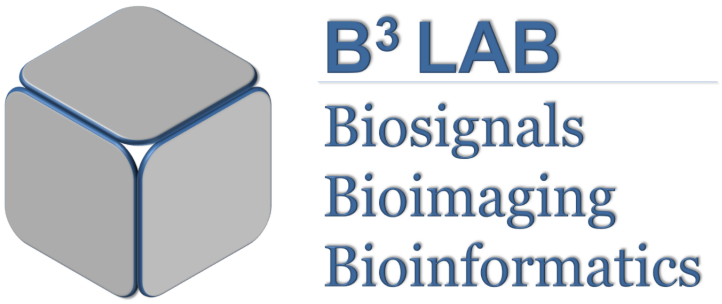
Thesis submitted to the
University of Coimbra for the degree of
Master in Biomedical Engineering

Supervisors:
Prof. Dr. César Teixeira (Universidade de Coimbra)
Prof. Dr. Anna Maria Bianchi (Politecnico di Milano)
PhD Sara Mariani (Brigham and Women's Hospital)

Milan, 2017

This work was developped in collaboration with:

B^3 Lab Politecnico di Milano



Brigham and Women's Hospital

Centro de Informática e Sistemas da Universidade de Coimbra

To my parents and my beloved family...

Agradecimentos

Este momento é o fim de 5 anos cheios de histórias e de amigos que levo para onde for. Apesar de nesta fase final ter partilhado o percurso apenas com um computador não seria possível a sua conclusão sem a contribuição de algumas pessoas. A vós vos agradeço.

Obrigada ao Professor César Teixeira por me ter concedido esta oportunidade de estar num projeto partilhado em Milão, e pela disponibilidade e experiência que partilhou comigo. Obrigada à Professora Anna Bianchi por me ter recebido no B^3 lab e pela crítica construtiva que esteve sempre presente. A todos os membros do B^3 lab , mas particularmente ao Eros e ao Lucca que me receberam de braços abertos.

Obrigada a PhD Sara Mariani pela dedicação, por ter sido incansável. Não posso deixar de agradecer pelas inúmeras horas passadas no Skype a projetar o que hoje é a minha tese. Por último, por me ter ensinado e me ter motivado a aprender sempre mais e mais.

À minha família de praxis, à Rita, à Sami, ao Gui e à Ana por me terem feito voltar a apaixonar por esta cidade que é Coimbra, nunca esquecerei estes "verdes anos" e todos os momentos que aqui passamos.

Aos meus amigos de faculdade que levarei além Coimbra. À Ana Monteiro, à Cândida, à Maria Inês, à Rita, à Verónica, e ao Rebelo por estarem aqui incondicionalmente ao longo destes cinco anos, pela partilha diária de escárnio e mal dizer, por aturem o meu mau humor e por todos os momentos de diversão que partilhamos juntos. À Louro e ao Eduardo por me terem acompanhado neste percurso e por serem pessoas extraordinárias. À Gloria e à Kleopatra por todos os momentos inesquecíveis passados em Milão, e pelas viagens que fizemos juntas.

À minha família por me ter sempre motivado e por acreditar em mim e no meu sucesso. À tia Alice e ao tio Alcides, pelas palavras poéticas que sempre tiveram para mim e por me ajudarem a concretizar os meus sonhos. Ao meu tio Paulo e à

Titi Madrinha por me transmitirem a paixão pela física e por terem guiado sempre na melhor direção. À Cinita a ao Toni pelo vosso "sorriso de encantar" e por me terem cantado "trovas e cantigas de embalar". Aos meus avós e a emprestada avó Rosinha por me terem mimado da melhor maneira. A todos os meus primos que são mais que muitos, mas em especial ao Afonso e ao Miguel que são como os meus irmãos mais velhos que sempre olharam por mim.

Aos meus pais por terem movido mundos e fundos para tornarem o meu futuro melhor, e por terem investido sem nunca hesitar na minha educação. Ao meu pai por ter sido sempre exigente fazendo de mim uma aluna dedicada. À minha mãe por me amolecido um pouco mais o meu coração que sempre disse ser de pedra.

A todos os que tornaram este percurso inesquecível e incomparável, um grande obrigada!

Acknowledgments

This moment is the end of 5 years full of stories and friends that I take to wherever I go. Although this final phase was only shared with a computer it would not be possible to conclude without the contribution of some people. And for that I thank you.

Thank you to Professor César Teixeira for giving me this opportunity to be on a shared project in Milan, and for the availability and experience he shared with me. Thank you to Professor Anna Bianchi for having received me in the B^3 lab and for the constructive criticism that was ever present. To all members of the lab, but particularly to Eros and Lucca who welcomed me with open arms.

Thank you to PhD Sara Mariana for her dedication, for being relentless. I will have to thank for the countless hours spent on Skype to design what is today my thesis. Lastly, for having taught me and motivated me to learn more and more.

To my family of praxis, Rita, Sami, Gui and Ana for making me fall back in love with this city that is Coimbra, I will never forget these "past years" and all the moments that have passed here.

To my university friends that I'll keep after Coimbra. Ana Monteiro, Cândida, Maria Inês, Rita, Verónica, and Rebelo for being here unconditionally throughout these five years, for the daily sharing of gossip, for putting up with my bad mood and at all times of fun we share together. To Louro and to Eduardo for being by my side over the years and for being extraordinary human beings . Gloria and Kleopatra for all the unforgettable moments passed in Milan and for the trips we have done together.

To my family for always motivating me and for believing in me and my success. To Aunt Alice and to Uncle Alcides, for the poetic words they always had for me and for helping me make my dreams come true.To my Uncle Paulo and to Titi Madrinha for transmitting my passion for physics and for always guiding me in the best direction.

To Cinita and to Toni for your “sorriso de encantar”” and for singing to me “Trovas e cantigas de embalar.” To my grandparents and the loaned grandmother Rosinha for spoiling me in the best way. To all my cousins that are a lot, but specially to Afonso and Miguel who are like my older brothers, are for always taking care of me.

To my parents for having moved the world to make my future better, and for having invested without ever hesitating in my education. To my father for being always demanding making me a dedicated student. To my mother because I softened a little bit more my heart that she always said to be of stone.

To all who have made this journey unforgettable and incomparable, a big thank you!

"The world is a hellish place and bad writing is destroying the quality of our suffering."

Tom Waits

Resumo

O sono é um estado de descanso que desempenha um papel essencial na vida de muitos seres vivos, inclusive humanos. A atividade do cérebro durante o sono aumenta, o cérebro é ativo e responde aos estímulos externos como um buffer com respostas rápidas e flexíveis, gerando fases de sono. Esta macroestrutura do sono descreve as diferentes fases do sono e é caracterizada com base nos diferentes ritmos do eletroencefalograma (EEG). Os eventos transitórios também são uma característica importante para caracterizar o EEG do sono.

Nesta tese, propõe-se um classificador robusto para a deteção desses eventos transitórios, especificamente arousals e complexos K. Os arousals são fenômenos periódicos que perturbam o sono e os complexos K são ondas padronizadas estereotipadas do EEG humano. O processo de classificação visual desses dois eventos é usado para inspecionar a qualidade e a fragmentação do sono e ajudar na classificação das fases do sono.

Para remover o ruído dos sinais, foram utilizados dois detetores de artefactos diferentes, o primeiro utilizou análise de entropia com a técnica Multisclale entropy e o segundo utilizou a análise espectral de potência do sinal de EEG. Em seguida, a remoção da interferência de ECG foi também aplicada aos sinais. Para detetar arousals, duas técnicas foram comparadas: técnica de Spectrograma e Multitaper. A detecção de complexos K foi testada usando filtros matched. Esses métodos foram validados em um conjunto de dados de 40 indivíduos de dois bancos de dados diferentes: MESA e MrOS da National Sleep Research Resource. Os algoritmos foram testados alcançando AUC de 0.804 para o espectrograma, 0.853 para a técnica Multitaper na classificação dos arousals. Para a deteção de complexos K, os filtros correspondentes foram testados apenas em um único sujeito que continha a classificação visual, obtendo um AUC de 0.814. A análise qualitativa da deteção do complexo K no conjunto de dados completo mostrou resultados encorajadores em termos de distribuição específica da fase do sono.

Palavras-chave: EEG em sono , eventos transientes, arousals, complexos K, detector de artefactos, Espectrograma, Multitaper, filtro Matched

Abstract

Sleep is an essential resting state that plays an essential role in the life of many living beings, including humans. The activity of the brain during sleep is increased, the brain is active and responds to external stimuli as a buffer with quick and flexible responses creating sleep stages. This macrostructure of sleep describes the different sleep stages and it is characterized based on the different rhythms of the electroencephalogram (EEG). Transient events are also an important feature to characterize sleep EEG.

In this thesis it is proposed a robust classifier for the detection of these transient events, specifically arousals and K-complexes. Arousals are periodic phenomena that disrupt sleep and K-complexes are a stereotyped pattern waves of the human EEG. The visual scoring of these two events is used to inspect both the quality and fragmentation of sleep and to aid in the scoring of sleep stages.

In order to remove the noise from the signals two different artifact detectors were used, the first applied multiscale entropy analysis and the second used the EEG power spectral analysis. Then a ECG interference removal was also applied to the signals. To detect arousals two techniques were compared: Spectrogram and Multitaper technique. The K-complexes detection was attempted using matched filters. These methods were validated on a dataset of 40 subjects from two different databases: MESA and MrOS from the National Sleep Research Resource. The algorithms were tested achieving AUC of 0.804 for the spectrogram, 0.853 for the Multitaper technique in the classification of the arousals. For the detection of K-complexes, the matched filters was tested only on a single subject with the visual scoring, obtaining an AUC of 0.814. Qualitative analysis of K complex detection on the full dataset showed encouraging results in terms of sleep stage-specific distribution.

Keywords: sleep EEG, transient events, arousals, K-complexes, artifact detectors, Spectrogram, Multitaper, Matched filters

List of Figures

List of Tables

Acronyms

AASM American Academy of Sleep Medicine

AUC Area Under the Curve

ANN Artifical Neural Network

ASDA American Society of Sleep Disorders

CAP Cyclic Alternating Pattern

CI Complex Index

DER Detection Error Rate.

DPSS Discrete Prolate Spheroidal Sequence

DTW Dynamic Time Warping

ECG Electrocardiogram

EEG Electroencephalogram

EMG Electromyogram

EOG Electrooculogram

FDR False Discovery Rate

FFT Fast Fourier Transform

FIR Finite Impulse Response

FPR False Positive Rate

FNR False Negative rate

GUI Graphical User Interface

LDA Linear Discriminant Analysis

MA Micro Arousal

MCA Morphological Component Analysis

MESA Multi-Ethic Study of Atherosclerosis

MrOS Osteoporotic Fractures in Men Study

MSE MultiScale Entropy

NLEO Non-Linear Energy operator

nREM non Rapid eye movement

NSRR National Sleep Research Resource

OSA Obstructive Sleep Apnea

PPV Positive Predicative Value

PSD Power Spectral Density

PSG Polysomnogram

REM Rapid eye movement

R&K Rechtschaffen and Kales

ROC Receiver Operating Characteristic

S1, S2, S3, S4 NREM sleep stages of incresing sleep depth according to R&K

SS Sensitivity

SP Specificity

STFT Short Time Fourier Transform

TEO Teager Energy Operator

TPR True Positive Rate

Contents

Chapter 1

Introduction

1.1 Context

This dissertation has its focus on the study of the transient events of sleep, these phenomena are the smallest elements that compose the sleep, they are the core of sleep and complement the conventional approach of defining the sleep stages. Moreover, the events are visible on the Electroencephalographic (EEG) signal, the signal used to develop this thesis.

The EEG provides important and unique information about the sleeping brain. This signal has been used in medicine since the 20th century, starting for the first time when Richard Capton placed electrodes over the human scalp and tried to read the electrical signals of the brain [6] allowing to investigate the bioelectrical activity characteristic of the sleeping human brain. The popular myth that the brain activity was shut down during sleep was then crushed, actually the activity of the brain is strong and complex during sleep.

The quality of sleep is essential for a good cerebral activity. During sleep the EEG shows changes within the preponderant waves and the presence of the transient events studied in this work, is due to a complex mixture of neurophysiological processes, that interact mutually to create a layered, efficient structure. The sleep process is characterized by different stages with different characteristics,(Macrostructure) and the Microstructure that defines the little transient events, both of the structures will be described in detail in Chapter 2.

The trasient events, have an essential role for the evaluation of the quality of the sleep. Arousals are associated with body movements, with respiratory events and if their appearance is frequent the sleep quality is decreased. K-complexes are a key factor for the definition of the sleep stages. Spindles are strongly connected to the memory formation and sleep diseases. The detailed definition of these events is described in the section 2.2.1, section 2.2.2 and section 2.2.3, respectively.

This project was developed in a partnership between University of Coimbra and Politecnico di Milano, the work was developed in the B^3 lab in Milan. Moreover another key factor for the development of this thesis was the collaboration and guidance of the researcher Sara Mariani at Brigham and Women's Hospital, in Boston, MA, USA.

1.2 Motivation

There is an ever-increasing global demand for clinical and healthcare services that are effective and affordable for the overall population. Biomedical signals are an important source of information in the medical world but their interpretation relies on the opinion of the experts analysing them. The EEG is a complex signal that for most of the physicians presents a challenge on its interpretation, besides even for an expert it is a very time consuming process. In fact, the manual classification of the EEG presents some reliability problems mainly because it is a subjective process [7].

Automatic detection of transient events on the EEG has been the subject of many studies, however, it is still considered not robust enough to replace the visual scoring. In this project we aim to develop robust classifiers for arousal and K-complexes, that are tested and compared with previous work on the subject.

1.3 Goals

The main goal of this project is the development of robust classifiers for the detection of two transient events, sleep arousals and K-complexes, using only a EEG signal. The software developed will be optimized to reduce the duration of the classification processes. The results will be compared with the previous work.

1.4 Thesis Outline

This dissertation begins with a brief contextualization of the theme of this thesis, followed by the motivation beneath this study, the main goals to be achieved and the outline of the written document. This is the Chapter 1.

Chapter 2 describes the sleep macrostructure and microstructure, focusing on the transient events that compose the sleep. The signals used and their characteristics are also presented.

Chapter 3 describes an overview of the previous works developed for the classification of arousals as well as K-complexes.

Chapter 4 details the materials used in this project, the databases and its corresponding signals. Moreover, it is described all the methods used for the treatment of these data and the detectors applied to them in order to classify the transient events.

Chapter 5 presents all the results achieved with this thesis and the comparison between them and the previous literature. As well as discussion of the validity and the importance of the results.

Chapter 6 recalls the initial project goals and compares it with the project output, also compares the classifiers with each other and with the results achieved in the literature.

Last, in Chapter 7 the conclusions obtained through the development of this project are presented, regarding the classifiers developed and their applications. Also it is mentioned all the future work that can be done in the area of study to improve the detections and in this specific study to produce a commercial interface.

Chapter 2

Sleep Concepts and Definitions

Sleep is a rest stage, observed in Humans and animals. It is an essential physiological phenomenon that takes part in the life of most creatures, [3] .When sleeping the brain is active and responsive to external stimuli although it is in an unconsciousness state. The arousal systems in the organism are responsible for controlling the duration and the depth of sleep. Therefore, it reflects the state of the organism and has the capability of affecting the mental and physical health of an individual. Different activation forces act upon the brain at the same time. The stimuli are hierarchically organized and the brain acts as a buffer with quick and flexible response to them, creating different stages within the sleep period [8] .

2.1 Macrostructure of sleep

It was in 1968, that for the first time, a set of rules were defined to score sleep in normal adults, established by a committee of experts chaired by Rechtschaffen and Kales [9]. The information used for the classification was extracted from polysomno-

graphic recordings (PSG) which includes multiple data entries : electroencephalogram (EEG), regarding the cerebral activity ; electromyogram (EMG), information from the muscle tone ; electrooculogram (EOG), eye movement data ; airflow, respiratory effort cardiac rhythm ; oxygen saturation and leg movements. The sleep recording was then split into smaller pieces , epochs of 30 or 20 seconds depending on the preferences and recording techniques of the sleep laboratory. R&K identified five different sleep stages, rapid eye movement (REM) and four non rapid eye movement (n-REM) stages (S1, S2, S3, S4) . Each stage is characterized by different prevalence of specific cerebral waves accompanied with physiological changes.

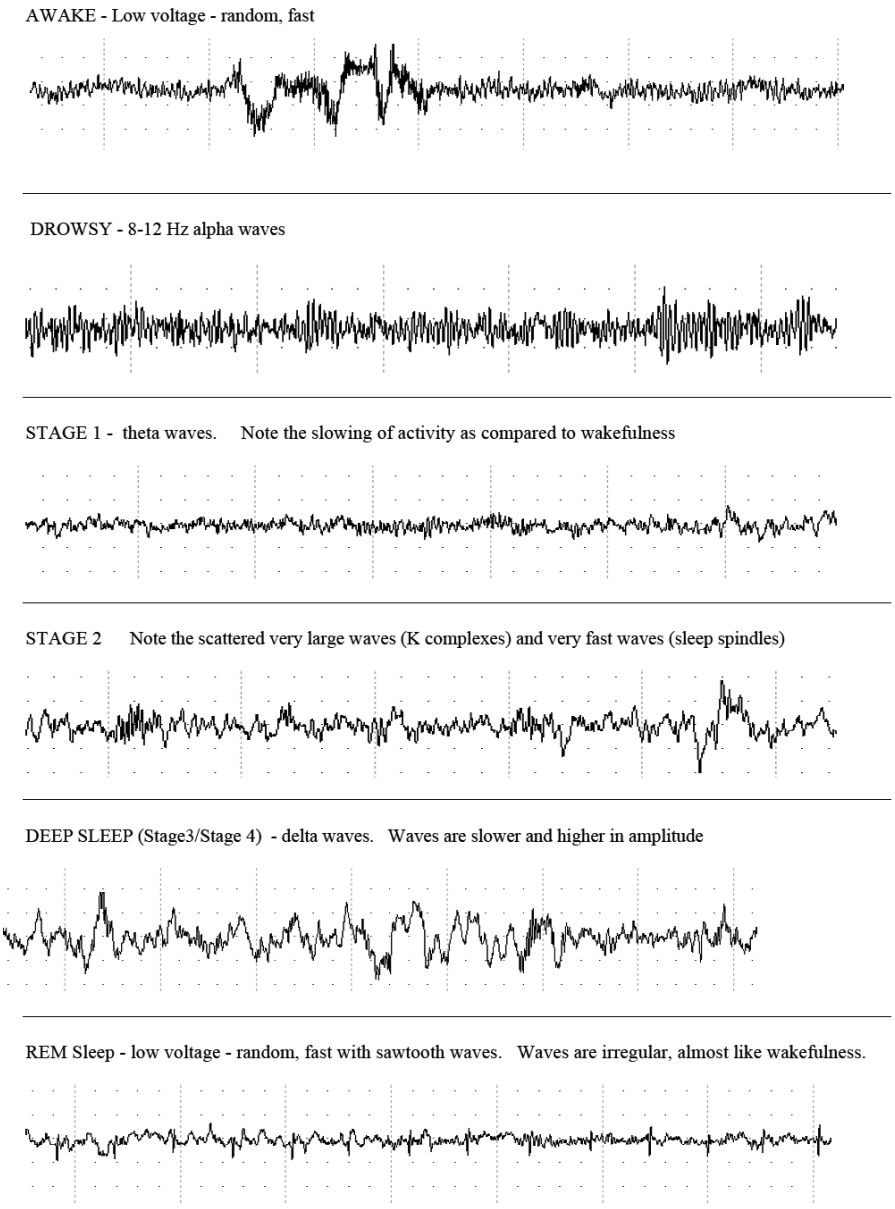


Figure 2.1: The different Sleep stages and the waves that compose them.(Adapted from: [10])

With these rules it was possible to get the same classifications in different labs and unify the classification of sleep but it neglected short-lasting events. In addition, this approach performs the classification in fixed size epochs which introduces flaws. So in 2007 this classification was revised by the American Academy of Sleep Medicine and some changes were implemented, [11]. The slow wave sleep stages S4 and S3 were fused in one reducing to 3 NREM sleep stages.

The EEG signal is composed by four different wave types, with different frequency ranges, from low to high frequencies, and are respectively: delta, theta, alpha and beta [11]. The properties of these waves define the sleep stage. For the quiet wakefulness stage the EEG signal has low amplitude, with at least 50% of the epoch composed with alpha rhythms. N1 is a stage of light sleep, it is the transition between sleep and awake. Through the life of the individual this stage suffers changes due to the the aging process but in adults the signal has a frequency range of 2-7 Hz with a presence of alpha waves in less than half the cycle. As the waves slow down and the low frequencies become dominant we change the Stage of sleep to N2. This stage is characterized by occasional burst of rapid waves, the presence of K-complexes and sleep spindles. In the deep sleep-stage, N3, delta waves begin to appear these waves have high amplitude and slow frequency. In this last stage sleep spindles still appear but with less frequency. The REM sleep is the part of sleep associated with dreams and there is a notable movement of the eyes , a rise on the heart rate and on the breathing rhythm and also an increase of the blood pressure. The EEG from this stage presents waves with low amplitude and mixed frequencies with a saw tooth like shape [3].

A number of different factors can affect the macrostructure of sleep, factors inherent to the subject like age or sleep disorders and external factors from the environment. Age is one of the factors that most affect the overall sleep structure [12, 13, 14], mental diseases are often associated with sleep disorders [15], for example narcolepsy which is a disabling sleep disorder. Studies also show that anxiety can create a peaceless sleep, with nightmares and subjects show a clear difficulty to wake up [16]. The Macrostructure of Sleep , defined previously, has been studied over the years to inspect the sleep quality, [17] but is inside the Microstructure of Sleep, that is, the transient events detected in the PSG during sleep that this thesis focus on.

2.2 Microstructure of Sleep

The concept of microstructure of sleep allowed to define events that lasted less than a normal epoch, i.e., less than 20-30 seconds. These events are not a part of a plateau in a sleep stage but were transient events, shifts, occurring during the change of stage. These events are arousals defined in the section 2.2.1, K-complexes in the Section 2.2.2 and sleep spindles in the Section 2.2.3.

2.2.1 Arousals

It was in 1992 that the first definition of Arousal was proposed by the the American Association of Sleep Disorders (ASDA) [18], as a periodic phenomenon that disrupted sleep. When the interruption caused by the arousal is not reversible the subject is now in a awakening stage. Arousals are defined as abrupt changes of EEG frequency content lasting at least 3 seconds and followed by at least 10 seconds of sleep. The frequency change is towards faster rhythms, usually above 16 Hz , has it can be seen in the figure 2.2 . It is also stated by the ASDA as one of the defining rules of arousal that those are not spindles.

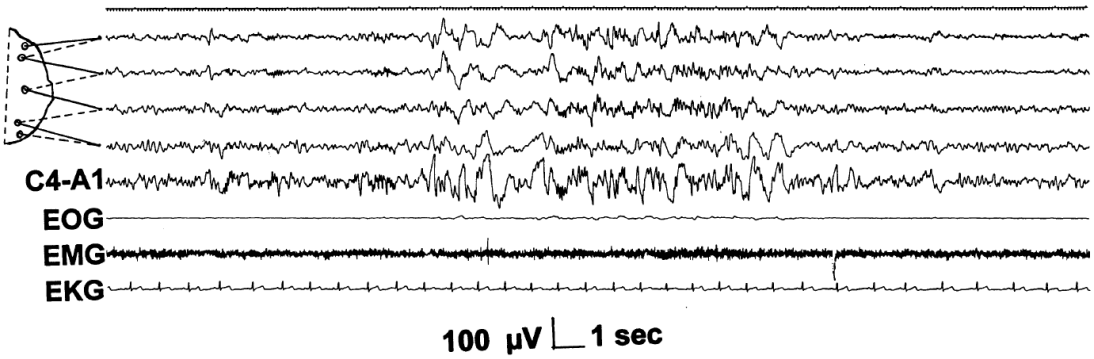


Figure 2.2: Example of an Arousal followed and preceded by sleep. (Adapted from: [19])

Arousals are mostly coincident with other physiological events for example respiratory events (apnea and hypopnea) [20] and body movements. The movement caused in an arousal is usually of the limbs, with some transitory intensification of the muscle tone and changes of the body position. Furthermore, a temporary increase of the heart rate was the third evidence of arousals. Arousals can be associated with other physiological events, as stated before including obstructive sleep apnea (OSA), and in this case the arousal is caused to keep the human being alive. For the subjects

with OSA, arousal can occur more than once per minute, deteriorating the quality of sleep significantly. Although arousals are normal and intrinsic components of sleep the amount of arousals define if the sleep is disturbed or not.

Drinnan et al. [21] conducted a study on EEG changes that accompanied the arousals and concluded that "EEG changes that accompany arousal can be quantified using a number of automated indices. All of the indices investigated show changes consistent with an increase in alpha activity or an increase in EEG frequency". The number of arousals is affected by aging, actually there is an increase of the number of arousals from the younger to the older subjects [13].

On the other hand, in REM sleep arousals were marked by a transitory lack of eye movements and alpha activities. In addition, these periods had a inconstant duration that could last from few up to 10 seconds. After the temporary activation a deactivation phase occurred generating a bi-phasic phenomenon. For a long period of time there was no interest in the deactivation phase and only Terzano and Parrino gave it significance when defining CAP in 1985 [22, 19] CAP is vital for the description of sleep quality because it describes with detail the dynamics of the brain during the night.

Due to the fact that arousals appear with other physiological events, in order to visually detect them it is easier to use a combination of signals, including, heart rate , EMG and respiratory signals. In REM sleep there are a lot of frequency changes, so it is necessary to use EEG data alongside with data from other biosignals. The EMG from the chin is usually selected to help in the classification of the arousals. The presence of the arousals in sleep is more likely to occur in superficial sleep stages, so more frequent in REM sleep and NREM stage 1.

2.2.2 K-complexes

A K-complex is an EEG pattern composed of three main events . The first is a negative peak of small amplitude, the second fast positive component with parabolic shape followed by a slower negative peak that also has an overall parabolic shape, the waveform is shown in the figure 2.3. The duration of this waveform can range from 0.5 up to 2 seconds, [23, 24]. K-complexes are a transient marker of NREM sleep appearing particularly in S2 but also in S3, so they are a feature used to score these sleep stages.

Moreover, these markers can be spontaneous as a marker of NREM sleep but also

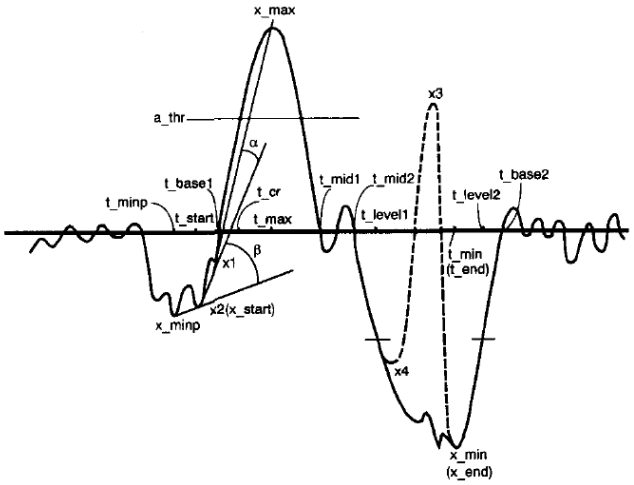


Figure 2.3: The approximate waveform representation of the K-complex. (Adapted from: [1])

may be the result of external auditory or somatosensory stimuli [25]. K-complexes can also be associated with other physiological events like vasoconstriction, an increase in the activity of the sympathetic system and in the arterial pressure. Due to these facts, K-complexes reflect the quality of NREM sleep. It is rare to find a K-complex by itself and it is usually associated with other rhythms alpha, delta or spindle. After the complexes long lasting changes occur to the EEG signal frequency and amplitude. Vertex waves are from the neurophysiological perspective, evoked potentials. In addition, these complexes and vertex waves result from the same cellular behavior but the difference between them is the phase of the sleep in which they appear, [26].

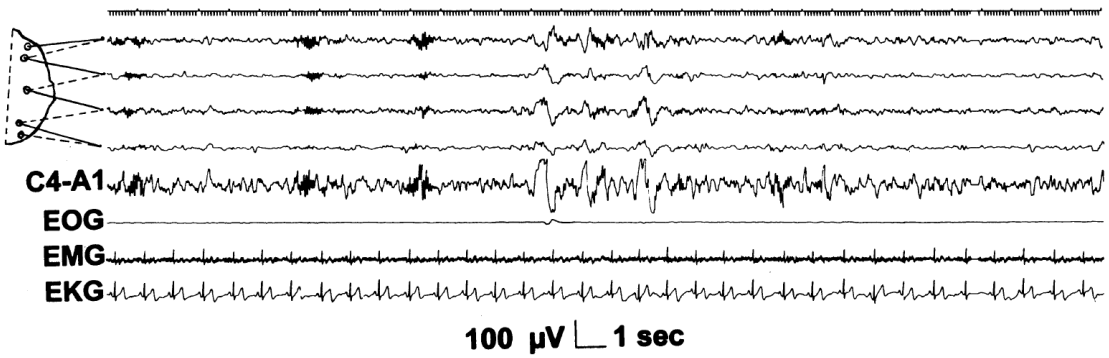


Figure 2.4: Example of K-complex sequences in S2 sleep. (Adapted from: [19])

2.2.3 Spindles

R&K were the first to define spindles [9] but over the years the definition evolved, there is still no exact definition but a spindle is a discrete event observed in the EEG signal. They are short lasting and powerful bursts of frequency between 11-16 hz with a waxing and waning shape that resemble an actual spindle, as it can be seen in Figure 2.5. They result from the powerful burst of synchronous neurons in the thalamus-cortical networks, and they define the S2 sleep stage. The duration, of the spindle, is of at least 0.5 seconds [18]. The changes in frequency within the spindle are connected to depolarization of thalamus-cortical neurons, [27].

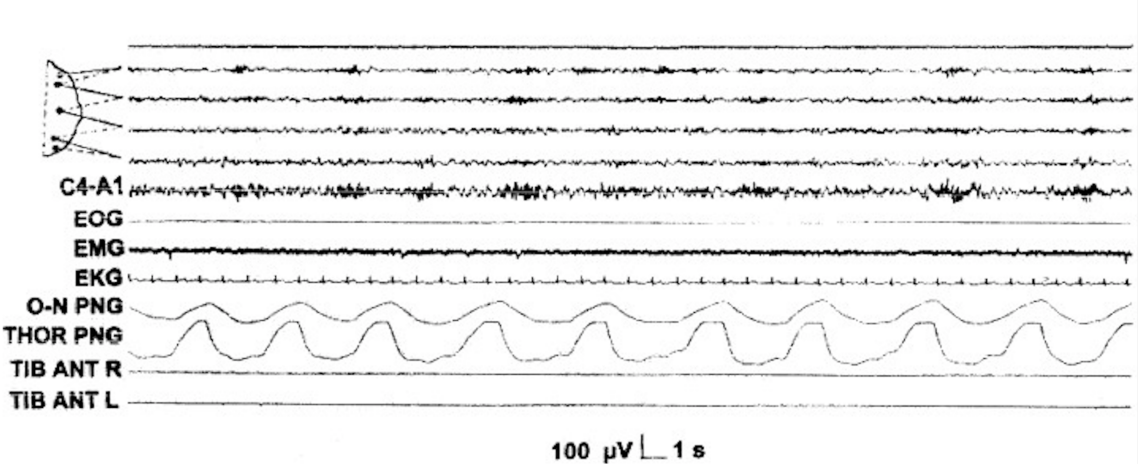


Figure 2.5: An example of mutiple Spindles in the EEG signal . (Adapted from:[28])

Initially, spindles were used to aid sleep staging as all the others sleep events defined so far, but over the years many studies show that spindles play a main role in memory formation and consolidation during sleep, also they are an indicator of the brain function and the sleep stability, [29, 30, 31]. The detection of spindles may be performed on different EEG leads, but it was in the frontal derivations that their presence showed a higher relevance to memory formation.

Moreover, spindles are associated with different pathologies including epilepsy and schizophrenia [32, 33]. For example, before epileptic seizures studies have shown a reduction of the power and density of spindles detected in the EEG signal. In the case of schizophrenia also the number of spindles is reduced dramatically and they appear to be less coherent.

2.3 PSG

In the late 1920's and 1930's PSG blossomed in the medical world in the fields of physiology, psychology and psychiatry [34]. Over the next 5 decades many studies occurred with the goal to find what was the most likely stage of sleep associated with dreaming. Currently PSG is essential to figure out if the sleep of an individual is normal and to track and monitor many sleep disorders [35].

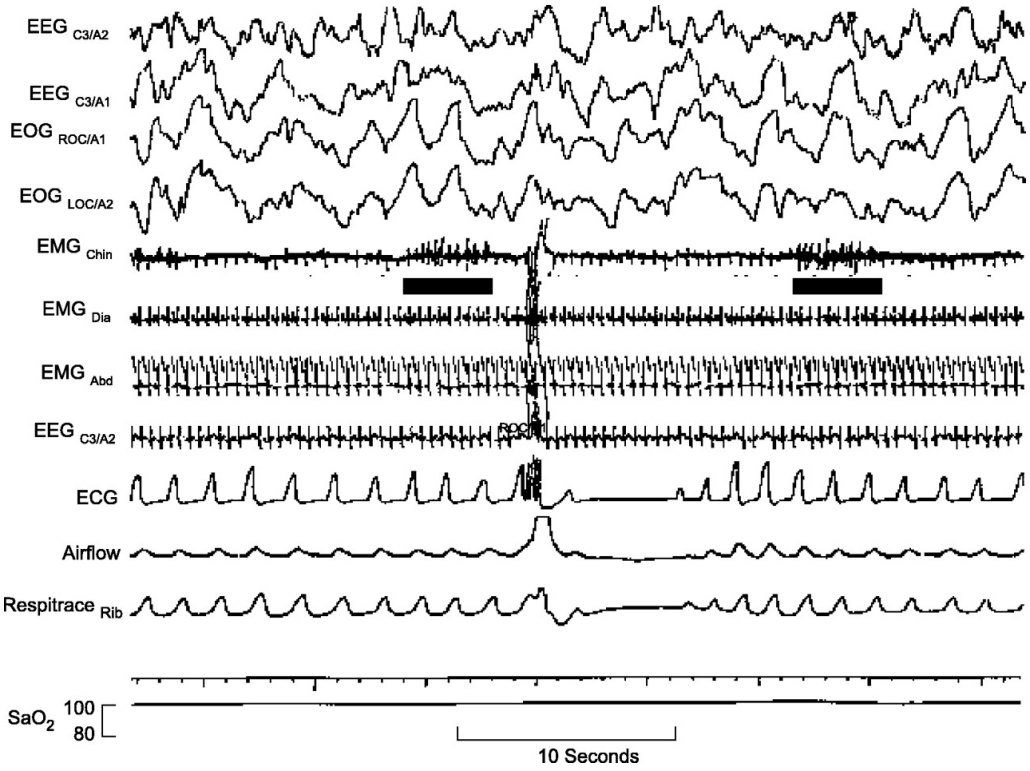


Figure 2.6: The Polysomnography of a normal individual, with all the corresponding channels. (Adapted from: [2])

The PSG is now a non-invasive pain free diagnostic procedure in which an individual must be asleep and many sensors are fixated to his/her body. It allows the movement of the person during the sleep and it must be done in a controlled environment. It was designed to record various physiological processes, for example the heart beat (ECG), the electric brain activity (EEG), the muscular activity (EMG), the movement of the eyes (EOG) . The sensors used are electrodes applied to the skin of the individual in specific body parts.

2.3.1 EEG

The Human brain consists of an intricate network of neurons and the number of elements of this network is in the order of 10^{11}. The brain is like a processor, that takes into account the information from raw data and transforms it into cognitions, storages it in the memory, and finds complex solutions to advanced problems.

To transit the information between cells, the neurons are activated and local current flows are produced. Electric potentials are formed between the synapses of axons and dendrites. Those currents are created of ions that are pumped through channels following the direction of the membrane potential.

The EEG allows to view the combination of the electrical activities of the neurons. To measure this electric activity many electrodes are displayed in a person's scalp in standard positions, as shown in figure 2.7. As for any electric measure a reference is needed, so a monopolar channel is defined as a signal acquired from a specific channel and a fixed reference.

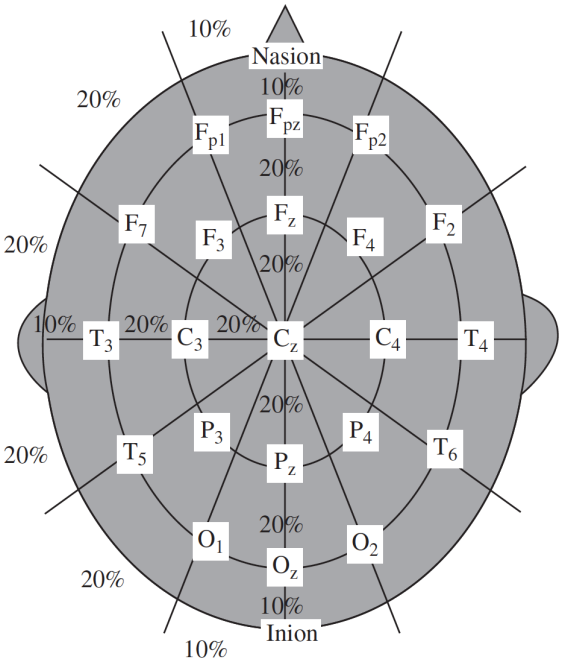


Figure 2.7: The international 10-20 disposition of electrodes placed over the scalp. (Adapted from:[3])

There is a significant noise component in the monopolar channels, due to the fact that they are taken at the body surface. It is possible to reduce these interferences by using bipolar channels. These are the difference between two monopolar channels.

Moreover the interference of other biological signals like EMG, ECG or EOG is less significant and they provide better spatial resolution when compared to the monopolar records [36].

2.3.2 ECG

The heart is the pump of the body, making sure that to every part of the system gets the right amount of blood. This cardiovascular event is required to sustain the life of the humans. The blood flow must follow a pattern so is necessary to maintain a regular cycle of relaxation and contraction.

The myocardium has specialized working cells that produce electrical impulses throughout the myocardium and they regulate the contraction of the heart. The electrical signals can be monitored with electrodes over time, and this measure is what is called the ECG [37].

In this thesis the ECG signal will be used to remove its interference with the EEG.

Chapter 3

Automatic Methods for the detection of Transient Events on Sleep EEG

Most of all transient events detectors share general steps, like pre-processing, feature extraction, feature selection, classification and performance evaluation. This chapter briefly describes along its sections the previous works of other authors on the previously enumerated steps. The automatic detection of transient events is typically based on time-frequency analysis or recognition of patterns on the EEG, potentially combined with the analysis of other signals.

The usual steps that make part of EEG-based detectors are :

1. *Pre-processing* - The idea of this step is to minimize the effect of noise and artifacts.
2. *Segmentation* - Consists of the division of the EEG into epochs by applying windows, that can be of fixed length or vary, depending on the local char-

acteristics of the signal. Adaptive segmentation methods like spectral error measure, nonlinear energy operator [38, 39], break up the signals into quasi-stationary segments of variable length.

3. *Feature extraction* - The appropriate feature set depends on the transient event to detect. For arousals an example is the power in the alpha and beta bands. For K-complexes the shape and the frequency range are in the core of the features extracted.

4. *Classification* - Many techniques allow a possible detection of the micro event. Examples are rule based classifiers using mainly thresholds and Machine learning techniques.

An alternative option to points 2 and 3 is the use of time-scale methods, like the wavelet transform, to obtain the spectral power in the frequency bands of interest with variable time-frequency resolution [40]. Moreover, in the case of arousals to detect them the information can be extracted not only from the EEG signal but from a combination of signals, ECG, EMG and respiratory signals. This is due to the fact that arousals are associated with other physiological like movements or obstructive sleep apnea. The suggested EEG leads for this type of analysis are a central and an occipital lead, to facilitate detection of beta and alpha activity, respectively.

3.1 Detection of Arousals

Arousals are transient events of variable length therefore an adaptive segmentation is a very suitable technique. In adaptive segmentation the boundaries between segments are set when significant changes are detected [41].

In 1999, Zamora et al. [42] used Artificial Neural Networks (ANNs) to analyze the EEG signal to detect micro-arousals [43]. ANN is a interconnected network that mimics the complexity of the brain and its decision process. The features were extracted with a bank of band pass filters and with an autoregressive method to estimate the power spectrum. From 88% up to 100% of the arousals were correctly classified with this technique. The dataset included only the epochs in which the 3 experts agreed on the visual classification.

Carli et al. [44] also used a Machine learning technique to detect the arousals. In this study the information from the EEG is combined with the information from the EMG. Wavelets were used to compute the power spectrum in different bands

and time frequency analysis was used to evaluate groups of indices describing the changes on the EEG [45, 46]. These indices represent the ratios between short and long term averages of each band of frequencies. A Linear Discriminant Analysis (LDA) [47] is then used to classify the data. The score of the two EEG channels (F4-C4 and C4-O2) and the EMG channel is then weighted and a final classification is achieved. A sensitivity of 88.1% and a precision of 74.5% were achieved.

Sugi et al. [48] combined the information from four EEG channels (C3-A2, C4-A1, O1-A2 and O2-A1) and the respiratory information (TFLO and CFLO). The features were extracted from the spectral density achieved with fast Fourier transform (FFT) applied to 1-second epochs. These are mainly parameters of duration, frequency and amplitude of the delta and alpha waves. The results were then limited by a threshold to determine the arousals, and achieving an accuracy of 77%.

In 2005 Agarwal et al. [49] developed an automatic classifier of MAs. First, for preprocessing an adaptive noise canceller filter was used to remove the ECG interference in the EEG. For the segmentation step, a adaptive segmentation technique was used, creating segments with a minimum size of three seconds. Frequency-weighted energy defined by the non-linear energy operator (NLEO) [39] was used to enhance the alpha and beta bursts in the segmentation scheme. The features extracted were five, derived from the power bands and the maximum amplitude at each seconds. Next, the authors employed rule-based selection to ignore candidate arousal that did not respect the ASDA rules including the 10 seconds pause between arousals. Sensitivity ranged from 70.1% to 82.2% and specificity from 56.6% up to 72.4%.

Cho et al. [50] retrieved the information from the bipolar channel C3-A2 of the EEG to classify arousals. In the preprocessing several artifact removers were used, including a band pass filter was applied to the EEG with a frequency range from 0.5 to 50 Hz to remove noise. A spectrogram with 257 points with a Hanning window was applied to the signal and a power spectrum was computed every 60 seconds. Six bands of frequencies and the ratios between them represented the features. As a classifier an SVM (support vector Machine) with a radial basis function as kernel was used. Finally, the arousal with a duration shorter than 3 seconds were rejected. A Sensitivity of 75.26% and a specificity of 93.08%, were reported.

Shimel et al. [51] introduced a new technique that mimicked human behaviour of learning and detecting. First they defined the critical points, points in time in which significant change occur in the two EEG signals from the bipolar channels C4-A1 and C3-A2. Then, a classifier that applies a data mining technique based on the

Sequential pattern discovery was applied [52, 53]. The approach needs the data to be event oriented rather then having the data split, for example, in fixed size epochs. Finally it uses meta-rules and actual-rules to classify the data. Meta-rules are rules that governs the application of the actual rules. Using the data only from the EEG it achieved and average sensitivity of %49.4 and a positive predictive value (PPV) of %82.5 , if the data from the EMG is combined the average sensitivity rises to 63% and the PPV decreases to 82%.

In 2009 Sugi et al. [54] combined the data from four EEG channels(C3-A2, C4-A1,O1-A2 and O2-A1) with the data from the airflow and EMG. The data is segmented in blocks with fixed size and the periodograms were calculated using FFT with a Hanning window. The features extracted from the EEG consist of the power in the different bands and on an average of the results from the four channels. Finally for the detection it uses a threshold technique combined with duration rules to ensure the AASM definition of arousal. The average sensitivity, false positive rate (FPR) and FNR achieved were of 96%, 12% and 18%, respectively.

Behera et al. [55] used ANNs to classify this transient event. This study also combines the data from multiple channels of the PSG: two EEG bipolar derivations (C4-M1 and C3-M2) and an EMG channel. In the preprocessing stage it uses bandpass filters to extract the alpha and the beta frequency bands (8-30Hz). This method detects an event when the Power spectral density crosses a threshold and it also combines the information from the Hjorth parameters. From each event are extracted features from the power in all the frequency bands (Alpha, Beta,Theta, Delta and sigma). Finally an ANN is feed with all the features from the events, achieving an average specificity of 91.4%, an average precision of 91.7% and an average area under the receiver operating characteristic (ROC) curve (AUC) of 92.3%

Liang et al. [56] introduced a new SVM technique called Curious Extreme Learning Machine which“ is a fast, easy to implement machine learning algorithm based on a single hidden layer feedforward neural network.” In this work the data was extracted from the two EEG channels C4-A1 and C3-A2. Then it was filtered with a low pass filter with a frequency range of 0-50Hz. The signal was segmented in one second epochs and 22 features were extracted with frequency analysis of each epoch and the ratio between each neighbour epochs. The curious extreme learning machine was then feed with all the information and it retrieved a average AUC of 0.85.

Wallant et al. [57] incorporated the data from the EEG and the data from EMG. The novelty in this article is the introduction of two artifact removals techniques that minimize the effects of the movement of the subject and the effects of technical

faults. First, the EEG data is filtered using a band-pass filter (0.5-35Hz) that has the lower cutoff frequency different from the usual aiming to remove the sweating and breathing interference. Second, the data is segmented in fixed time epochs with a duration of 1 second. For feature extraction it was used power spectral density (PSD). For the final detection it uses thresholds, and it achieves a sensitivity of 76% and a specificity of 76%.

The latest work developed for EEG arousal detection was developed by Fernandez et al. [58]. In the pre-processing phase the signal is notch filtered at a frequency of 50Hz to remove the power line interference. Then a short-time Fourier transform using a sliding Hamming Window with a size of 3 seconds and with a shifting step of 0.2 seconds was applied. The point is to extract the "power content at a certain instant of time with the corresponding baseline levels from the immediate past instants". The information from EEG is combined with the ECG signal. After the events are marked the processing is focused only on them, comparing them with the spectral power of past ones and analyzing their shape. An innovative aspect is the analysis of the signal amplitude. A rejection step is the last, aiming to remove the arousals that do not follow the AASM definition or that do not have an EMG activation. This work achieved a precision value of 86% .

In conclusion, taking into consideration the literature presented above this thesis used, for the detection of the arousals, two techniques of time and frequency analysis, very common in the literature above to estimate the power spectrum of the signal. The novelty introduced is a specific technique called Multitaper.

3.2 Detection of K complexes

K-complexes are one of the key features to determine sleep stages and therefore over the years several investigators have done research regarding the best methods to detect them. In 1970, Bremer G. et al. developed the first K-complex classifier. This first project used the signal from three EEG channels F1-F7, P1-T5, O3-OZPZ and a technique of matched filters [59, 60] to detect this transient event. The main goal was to replicate the process of human visual classification so the waveform was mimicked and then analogic and digital filters with thresholds were used. The main conclusion of this study is that, although this first technique is primitive "the results indicate that although the accuracy of the detection system is argumentative, it is nevertheless comparable to that of Human scorers", [4].

Over the years many other authors used the technique of matched filters but only in the digital domain.

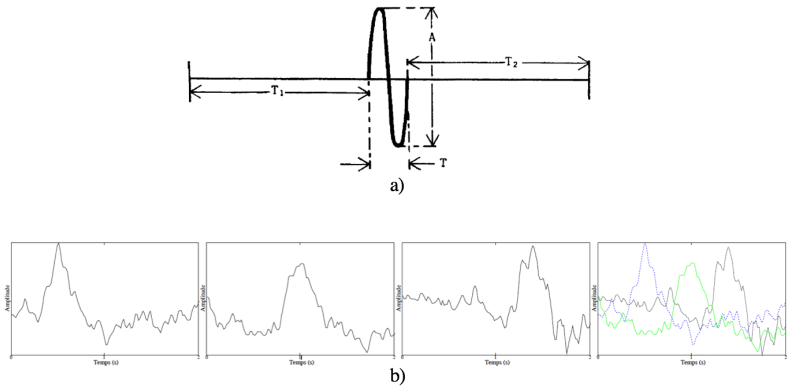


Figure 3.1: The evolution of the template of the matched filter. a) the template used by Bremer et al. [4] ; b) the three templates used by Kerkeni et al. [5]

In 2004, Woertz et al. [61] used the data from the bipolar channels C3–A2 and C4–A1 from EEG signal and compares two different methods. First, a time-domain analysis of the activity of low frequencies achieving a true positive rate (TPR) of 81%, a false positive rate of 7%. Second, applying matched filters using the same data the results increased both the TPR and the FPR, 93% and 32% respectively. In this study the filters used were based on the asymmetrical pattern of the waveform, using three different templates of different durations from 0.5 seconds up to 1.5 seconds. The final score results from the combination of three filters and then were classified accordingly to a threshold criterion. One particularity of this study is that the threshold is calculated based on the manual classification of 45 experts. Kerkeni et al. [5] also used matched filters but introducing a new concept, dynamic time warping DTW [62] which is an algorithm that allows time compressions and expansions on the templates of the waveform. The waveforms are three of different shapes and and durations. A TPR of 53% and an FPR of 10% were achieved in this study.

Other methodology used to classify this transient event was ANN's [43]. It was first used in 1990 by Jansen [63] in a multilayer ANN trained with backpropagation algorithm. Data was retrieved from the EEG channel Fp1 and 10 features were extracted considering the shape and the frequency range in the K-complex. A maximum recognition rate of 67% was achieved. Bankman et al. [1] tested the ANN with backpropagation fed with raw EEG data or with features. In [1] it was

also taken in consideration a more detailed evaluation of the shape of the K-complex. A TPR of 90% and a FPR of 8% were achieved using the features.

Moreover, Shete et al. [64] uses also ANN and compares them with the matched filter technique. For feature extraction they used the likelihood thresholds based on measures of amplitude and duration of the different phases of the K-complex. Shede and his team achieved a sensitivity of 96% and a specificity of 52% with neural networks and a sensitivity of 86% and a specificity of 67% using the matched filters. These features were also applied in EEG data by Devuyst et.al [65] but with fuzzy thresholds. The use of ANN in the study improved the results when compared to the work done with the same features with the fuzzy technique.

Rosa et al. [66] used an algorithm based on a stochastic model [67] that is "used to characterize the dynamics of the tonic and phasic activities of human sleep EEG". This model replicates the sleep with feedback loops of rhythms present in sleep (alpha, beta, delta, theta and sigma) with white noise and pulses corresponding to the K-complexes and the vertex waves. The detection algorithm works in the opposite direction to detect the transient events. The EEG derivation studied was C4 - A1 + A2. The average TPR and FPR were 82.4% and 53.2% respectively.

Another method used for the detection of the K-complexes was based on Wavelets [68]. This approach was first introduced by Tang et al. [69] in the detection of this transient events. Later, Erdamar et al. [70] used this method to classify the selected EEG data from the bipolar channel C3-A2. Then the signal was classified by two methods, wavelets on one side, Teager energy operator (TEO) on the other side [71]. Afterwards, the results were compared and a final label arose. An average TPR of 85% and an average FPR of 7% were achieved. Laerke Krohne et al. [72] pick up the previous work from Erdamar et al. and decided to improve it. First step was to change the Wavelets algorithm, the results were significantly better so the TEO implementation was no longer necessary and therefore it was discharged. Furthermore, 4 new features were introduced in a rejection step improving the PPV. The results were a TPR of 74% and a PPV of 65%.

In 2017, a research Group from Montreal created a computational application to detect both K-complexes and sleep spindles called Spinky [73, 74]. A Tunnel Q-factor wavelet transform [75] with a morphological component analysis (MCA) [76] was applied to decompose the signal collected by the C3 electrode into two components: an oscillatory and a transient one. The transient component is the important one given that it is associated with K-complexes. Then this component is filtered with a finite impulse response (FIR) filter and a threshold is applied to find the transient

events. This study achieved a sensitivity of 78.72% and a FDR (False discovery rate) of 23.44%.

Yazdani et al. introduced a novel algorithm for the extraction of short term events in EEG [77], including K-complexes. This method is a non-linear filter named Relative-Energy. The algorithm combines the information of the short and long term energies of the signal and provides a vector than enhances the points of interest. The channels used for this study were CZ-A1 and C3-A1. It was achieved a maximum PPV of 80.98% , a sensitivity of 65.05% and a detection error rate (DER) of 1.29%.

Many other techniques were tested over the years. Pohl et al. [78] introduced a Neuro-Fuzzy Detector combining Fuzzy logic with ANN. A joint Time and Time-Frequency detection was introduced by Richard et al. [79]. This study uses also an optimal method retrieved from the training data. Kam et al. presented a detector using a continuous density Hidden Markov Model [80]. In 2010 a study conducted at Instituto superior técnico (IST) uses a method based on Hjorth parameters and on fuzzy decision mechanism [81]. The machine learning technique was also used for the detection of this transient event [82]. In this paper also a new feature extraction technique is implemented.

From all the literature studied for the detection of K-complexes, Matched filters is a very common technique , that although was one of the first to be tested it is still used today, achieving statistically significant results. The shape of the K-complex is very particular and it is their essential feature, taking all this into consideration the detection of the K-complex was performed with matched filters.

Table 3.1: Summary of the methods present in the literature for arousal detection.

ref	*year*	*# of PSG recordings*	*channels*	*method*	*ACC*	*SS*	*SP*	*Precision*	*AUC*
[42]	1999	9	-	BandPass filters + ANN	-	0.55	-	0.88	-
[44]	1999	11	F4-C4 + C4-O2 +EMG	Wavelets + LDA	-	0.881	-	0.751	-
[48]	2003	8	C3-A2 + C4-A1 + O1-A2 + O2-A1 + airflow	TFA + Thresholds	0.77	-	-	-	-
[49]	2005	2	EEG+ECG	Adaptive SEG + Decision rules	-	0.822	0.724	-	-
[50]	2005	9	C3-A2	TFA + SVM	-	0.753	0.938	-	-
[51]	2009	26	C4-A1 + C3-A2	Meta+Decision rules	-	0.494	-	0.825	-
			C4-A1 + C3-A2 + EMG		-	0.63	-	0.82	-
[54]	2009	8	C3-A2 + C4-A1 + O1-A2 + O2-A1 + airflow + EMG	TFA + Thresholds + Rules	0.94	0.86	0.971	-	-
[55]	2014	10	C4-M1 +C3-M2 + EMG	Bandpass filters + PSD + ANN	-	-	0.914	0.917	0.923
[56]	2015	1	C4-A1 + C3-A2	C-ELM	0.79	-	-	-	0.85
[57]	2016	60	EEG+EMG	Artifact Detectors+ PSD + Thresholds	-	0.76	0.76	-	-
[58]	2017	22	EEG+ECG +EMG	TFA+ Decision Rules		0.75	0.99	0.86	0.87

Table 3.2: Summary of the methods present in the literature for K-complex detection.

ref	*year*	*# of PSG recordings*	*channels*	*method*	*SS*	*SP*	*Precision*
[4]	1970	-	F1-F7 + P1-T5 + O3-OZPZ	Matched Filter	-	-	-
[61]	2004	45	C3–A2 + C4–A1	TFA	0.81	0.93	-
				Matched Filter	0.93	0.68	-
[5]	2007	10	EEG	Matched Filter +DTW	0.53	0.90	-
[63]	1990	6	Fp1	ANN	0.67	-	-
[1]	1992	1	EEG	ANN	0.90	0.92	-
[64]	2012	-	EEG in nREM 2	Matched Filter	0.8647	0.6766	-
				ANN	0.9606	0.5262	-
[66]	1993	19	C4 - A1 + A2	Stochastic Model	0.824	0.468	-
[70]	2012	-	C3-A2	TEO + Wavelets	0.85	0.93	-
[72]	2014	10	Cz-A1	Wavelets	0.74	-	0.65
[73, 74]	2017		C3	Tunnel Q-factor wavelet + MCA	0.7812	-	0.7656

Chapter 4

Materials and Methods

4.1 Materials

The data analyzed in this thesis belongs to the NSRR, the American National Sleep Research Resource [83]. From many databases available two were carefully selected to be the subject of study in this thesis. From each database the first 20 subjects were selected.

4.1.1 MrOS

The Osteoporotic Fractures in Men (MrOS) Study was created to evaluate the effect of the fractures in the quality of life, in this specific case their effects on sleep [84]. Although the goals of this thesis are very distinct, the data available is ideal for the purpose of the study.

The subjects are 2911 men with age over 65 that were selected due to previous fractures [85]. Two sleep study cycles were made over the years, one from 2003 to

2005 and another from 2009 to 2015. For each visit of each subject it is available a polysomnographic recording of the night of sleep in a European Data Format (EDF) accompanied by a text file with the manual annotations of and the record of some macrostructure events, like arousals and respiratory events.

The subset used from this database is reported in Table 4.1 .

Table 4.1: MrOS Dataset used for the analysis and development of this classifier.

Subject	*Filename*	*Total sleep time (s)*	*Total sleep time (h:m:s)*	*Number of arousals*
1	mros-visit1-aa0001	806400000	8h45'00	54
2	mros-visit1-aa0003	7011840	7h36'30	74
3	mros-visit1-aa0005	11128320	12h04'30	105
4	mros-visit1-aa0006	9971456	10h49'11	87
5	mros-visit1-aa0009	12080640	13h06'30	175
6	mros-visit1-aa0012	7226880	7h50'30	106
7	mros-visit1-aa0013	7426560	8h03'30	83
8	mros-visit1-aa0014	9523200	10h20'00	99
9	mros-visit1-aa0015	11504640	12h29'00	133
10	mros-visit1-aa0018	10229760	11h06'00	182
11	mros-visit1-aa0019	10606080	11h30'00	170
12	mros-visit1-aa0021	10106880	10h58'00	173
13	mros-visit1-aa0024	10867200	11h47'30	238
14	mros-visit1-aa0025	4300800	4h40'00	74
15	mros-visit1-aa0026	9561600	10h22'30	186
16	mros-visit1-aa0027	7910400	8h35'00	193
17	mros-visit1-aa0028	10129920	10h59'00	164
18	mros-visit1-aa0029	11527680	12h30'30	273
19	mros-visit1-aa0030	10583040	11h29'00	253
20	mros-visit1-aa0031	8102400	8h47'30	85

The MrOS database includes PSGs with multiple channels, as stated before, including EEG and ECG. In the case of EEG there are 4 channels: two are bipolar channels (C3-A2 and C4-A1) and the other two are monopolar, A1 and A2. For this specific database the channel C4-A1 will be used. Data was acquired at a sampling frequency of 256 Hz and was high pass filtered at 0.15 Hz. The ECG information will be used to remove the interference in EEG.

4.1.2 MESA

The second database used is, Multi Ethic Study of Atherosclerosis (MESA). MESA was created to evaluate the differences in cardiovascular diseases and their progression in 4 different ethnicities, black, white, Hispanic, and Chinese-American. With a total of 2237 subjects with ages between 45 and 84 both men and women. Like the previous database this one also includes a PSG file correspondent to a night of sleep for each subject and a text file including all the manual details of the sleep classification by an expert.

In this thesis the diversity of the subjects is not essential but it is important to analyze the applicability of the classifier. Diversity was one of the main reasons to lead to the selection of this database. From all the subjects the first 20 were selected and are reported in Table 4.2.

Table 4.2: MESA Dataset used for the analysis and development of this classifier.

Subject	*Filename*	*Total sleep time (s)*	*Total sleep time (h:m:s)*	*Number of arousals*
1	mesa-sleep-0001	11058944	11h59'59	197
2	mesa-sleep-0002	10137344	10h59'59	169
3	mesa-sleep-0006	8294144	8h59'59	189
4	mesa-sleep-0010	9215744	9h59'59	97
5	mesa-sleep-0012	10964480	11h53h50	161
6	mesa-sleep-0014	12902144	13h59'59	89
7	mesa-sleep-0016	9215744	9h59'59	124
8	mesa-sleep-0021	8294144	8h59'59	264
9	mesa-sleep-0027	8294144	8h59'59	124
10	mesa-sleep-0028	8754944	9h29'59	194
11	mesa-sleep-0033	8294400	9h00'00	334
12	mesa-sleep-0035	8294144	8h59'59	89
13	mesa-sleep-0036	8294144	8h59'59	122
14	mesa-sleep-0038	8294144	8h59'59	43
15	mesa-sleep-0046	9215744	9h59'59	125
16	mesa-sleep-0048	9215744	9h59'59	145
17	mesa-sleep-0050	9215744	9h59'59	88
18	mesa-sleep-0052	11058944	11h59'59	115
19	mesa-sleep-0054	9676544	10h29'59	179
20	mesa-sleep-0056	11058944	11h59'59	116

The Mesa database includes PSG with multiple channels, including EEG and ECG. In the case of EEG there are 3 channels Fz-Cz, Cz-Oz and C4-M1. In this thesis the channel Cz/Oz will be used considering its favorable central position over the

occipital lobe. Data were acquired at a sample frequency of 256 Hz and low passes filter of 100 Hz. All the files used in this study were imported to *Matlab* using appropriate functions. The analysis of the signals and all the statistical work was done in the *Matlab*.

4.2 Methods

In this Section, the methodology is explained and the different steps are represented in the Figure 4.1. The methods used for the detection of the K complex differ drastically from the detection of arousals, for example the segmentation is not an essential step, the detection uses different methods and the classification is also distinct. In the next sections a more meticulous description of all the methods used in this thesis will be presented.

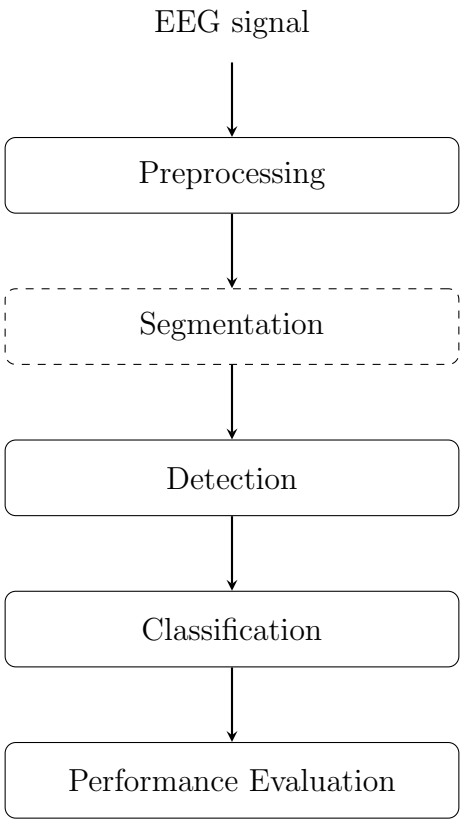


Figure 4.1: Diagram of the methodology used for the detection of the transient events.

The signals are retrieved from the EDF file with an auxiliary function created in Matlab to read these files. Moreover other function is used to open and read the annota-

tion files in XML format and to access the expert classification of the macrostructure of sleep, the hypnogram and the position and duration of the transient events.

4.2.1 Preprocessing

The preprocessing is essential to minimize all the undesirable noise and external interferences, for that reason multiple techniques are used to minimize them, as seen in Figure 4.2. Three methods are used, two complimentary algorithms to detect and remove artifacts and one algorithm remove ECG interference. As mentioned before, we employed different EEG leads from the two databases, so, although the steps are similar, the details of each step vary with the database. The signal provenient from the MESA is not contaminated with the heart beat therefore it is not necessary to apply the ECG removal.

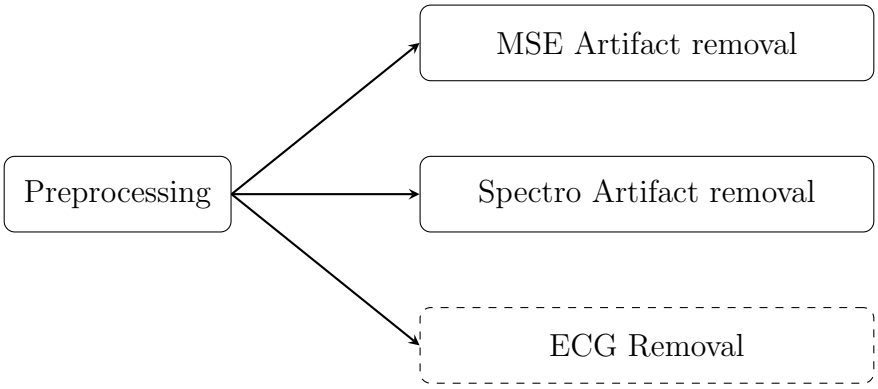


Figure 4.2: Diagram of the multiple preprocessing methods used.

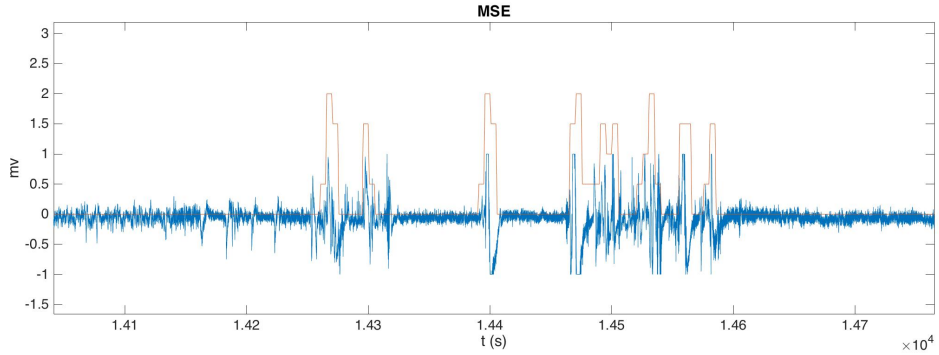
In the next sections a closer look into each of the steps will be presented.

4.2.1.1 Multiscale Entropy analysis for Artifact Detection

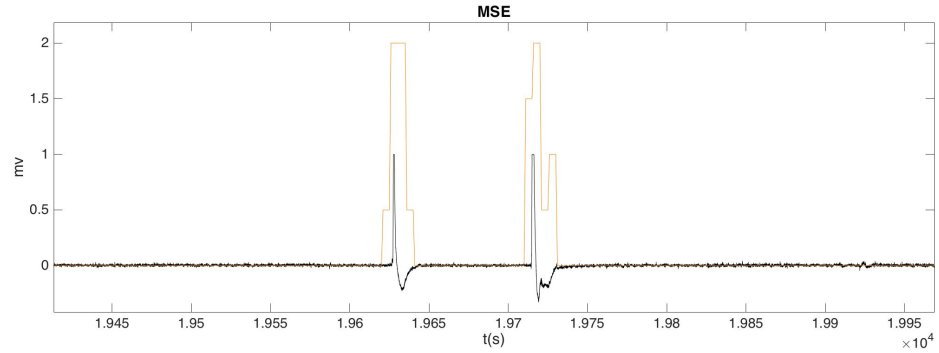
The goal of Multiscale Entropy (MSE) [86, 87] is to quantify the complexity of a signal. In the presence of an artifact the signal loses its information content and tends to incorporate either almost random processes (e.g. white noise) or very periodic events.

Entropy quantifies the degree of disorder and it is maximum for systems that are completely random. The Shannon entropy, is calculated by :

$$H(X) = -\sum_{x_i \in \Theta} p(x_i) log p(x_i) \tag{4.1}$$



(a)



(b)

Figure 4.3: The artifact scoring with the MSE technique in the EEG leads; (a) the results for a subject from the MrOS database ; (b) the results for a subject from the MESA database.

where X represents the variable with a maximum value of Θ and $p(x_i)$ is the probability mass function.

Due to the fact that the signals have a finite length that may be small and the calculus of entropy by the Shannon's Method is influenced by the length of the sample a different method to calculate Entropy called the Sample Entropy [88]. The Sample Entropy is mathematically the negative of the natural logarithm of the conditional probability of the sequences that are similar to each other within a tolerance r, will continue so, after one more point is added to both sequences. And its mathematical expression is :

$$S_E(m,r,n) = ln\frac{\sum_1^{N-m} n_i'^m}{\sum_1^{N-m} n_i'^{m+1}} \tag{4.2}$$

r represents the tolerance ratio, m the number of components and n the value of the signal.

Multiscale Entropy is an extension of the Sample Entropy method that allows to

look at a signal with different levels of detail. First, the MSE method divides the signal into non overlapping windows of fixed length τ. Second, the points inside each window are averaged, and the result is a consecutive coarse-grained time series, where each element is calculated by:

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, 1 \leq j \leq \frac{N}{\tau} \tag{4.3}$$

Last, it is calculated the entropy measure (S_E) for each coarse grained time series. To facilitate the detection of the artifact a concept called Complexity Index (CI) is introduced [89]. Considering epochs of 10 seconds with 50% overlap we calculated the sum of the entropy for scales 1 to 3. The maximum CI scale that can be used is the fraction of the number of elements of the epoch (10×256) and divided it by 750, because the sample entropy is largely independent of the number of elements for samples with a cardinal bigger than 750 [88].

Finally, we scored as artifact all the epochs where CI was lower than an empirically chosen threshold of 1.9. The score of the artifact for each time epoch is calculated considering the average of the neighbour epochs. In the Figure 4.3 is possible to see the detection of the artifacts using this MSE adapted method. The detector used is a modification of the publicly available tool [90].

4.2.1.2 EEG power spectra for Artifact Detection

To complement the previous detector another artifact detector was used in series with it. To detect the artifacts it uses the evaluation of the Spectral power of EEG

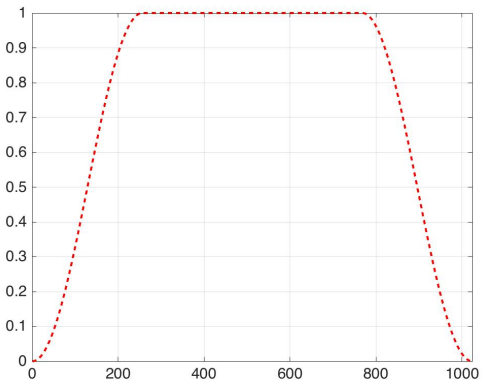
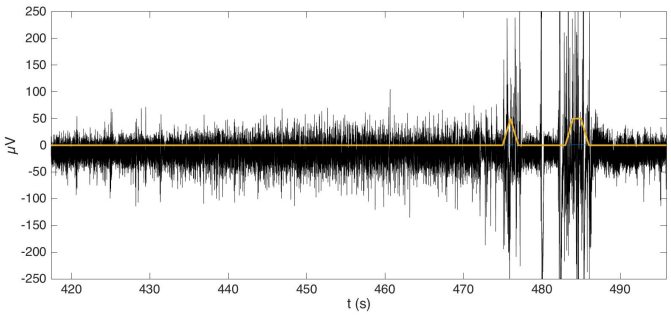


Figure 4.4: The Tukey window used for the calculus of the power spectra.

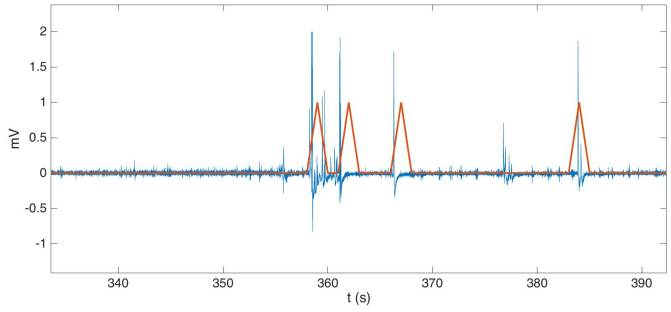
epochs. This has been used by other authors before [91], and to specifically detect artifacts [92].

First, it is computed the power spectrum of the slow frequency band [0.5-4.5Hz] and the power spectrum of the fast frequency band [20-40Hz] for consecutive 30 second epochs. The average spectrum is computed using the Welch's technique that averages spectra computed on overlapping sub-windows reducing this way the variance of the results. The window used for the calculation is the Tukey window with a dimension of 4s (4*256 points) and with a taper of 0.5 as seen in the Figure 4.4.

Second, it is computed a moving average encompassing 15 seconds of signal for both bands. Third, it is calculated for each power band the ratio between the current epoch and its respective moving average. For each epoch if the ratio is above a threshold it is considered to contain an artifact. The values of the thresholds used were 2.0 and 2.5 for the slow and the fast frequency bands. The final score is calculated by the sum of the results from each band, which means that if the interference is detected only in one of the bands it is scored as an artifact.



(a)



(b)

Figure 4.5: The artifact scoring with the Spectra power technique in the EEG leads; (a) the results for a subject from the MrOS database ; (b) the results for a subject from the MESA database.

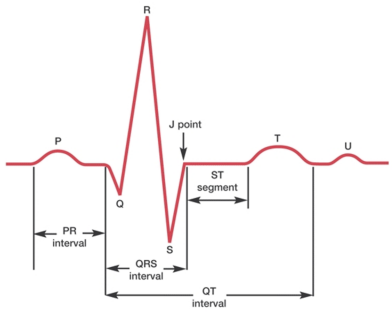


Figure 4.6: The shape of the ECG P-QRS-T complex. (Adapted from: [96])

4.2.1.3 ECG interference removal

Some of the EEG signals present a significant interference produced by the heart beat. In order to remove these and allow a better detection of the transient events a ECG interference removal was introduced.

Firstly, it was calculated the position where the possible peaks related to the heart beats and consequently related with the ECG peaks, were. It was used the Pan and Tomkins method to detect the QRS peaks [93] for three essential reasons: first, it is robust to the various types of interference that may appear on the ECG signal as a result of applying a Zero-phase forward filter [94, 95]; second, the use low thresholds that increase the sensibility of the detection, is allowed; third, the plasticity of the algorithm allows it to adapt to the EEG changes in shape and in frequency.

Secondly, the peaks that represent outliers, were removed. It was calculated the duration of the intervals between peaks, RR intervals. If the RR intervals were either too short or had a duration too different from the average duration of the neighbour intervals, they were eliminated.

Lastly, the interference of the ECG on the EEG was computed and removed. In order to do so, the EEG is divided into windows centered at each R peak previously detected. After, the average of all the windows is calculated. In case the ECG artifact is in the EEG the average window has a shape very similar to the shape of an ECG P-QRS-T complex, figure 4.6. Then, it was concatenate the average artifact n times, where n is the number of R peaks and the location of the peaks is used as guiding points. Finally, it was subtracted the artifact signal from the initial EEG and the clean signal is obtained, as it can be seen in the figure 4.7, [97].

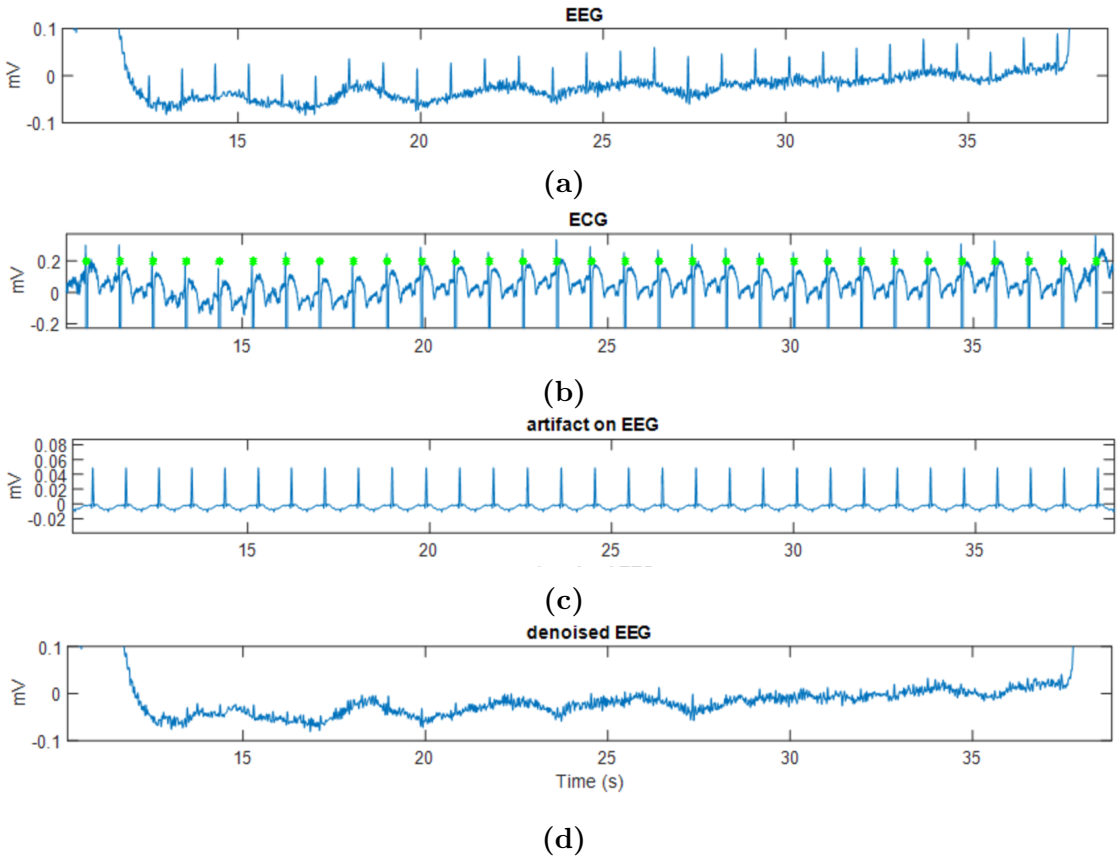


Figure 4.7: The ECG channel from the MESA database (a), the correspondent ECG signal (b) and the artifact it caused on the EEG channel (c). The result of the ECG interference removal technique on the EEG signal (d).

4.2.2 Arousal Detection

The spectral analysis has been recognized as a beneficial method for arousal detection since the early 80's [98, 99]. Taking into consideration the fact that arousals are acute changes on the frequency of the signal, two different methods of spectral analysis were performed, the traditional Spectrogram and MultiTaper and are presented in the sections below, 4.2.2.1 and 4.2.2.2, respectively.

4.2.2.1 Spectrogram

The spectrogram is a visual representation of the variation of the power of different frequency bands over time. In order to obtain this time-frequency discretization it is necessary to calculate the spectrum of frequencies for a specific stationary segment of a signal.

The Discrete Time Fourier Transform is the simplest form of frequency analysis.

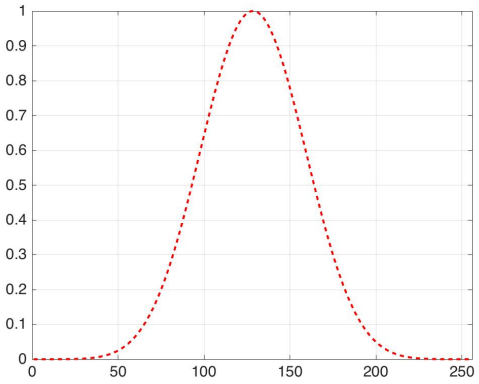


Figure 4.8: The Kaiser window used for the calculus of the power spectra.

But since the biomedical signals are not stationary and the goal is to study the changes of the spectrum over time it is necessary to study the signal as a group of segments and apply the Short Time Fourier Transform (STFT) [100], represented by the mathematical expression:

$$X\left[n,k\right] = \sum_{m=-\infty}^{\infty} x\left[m\right] w\left[n-m\right] e^{-j\frac{2\pi km}{N}} \tag{4.4}$$

where N is the total number of samples in a specific segment , $x[m]$ is the value of the signal at instant m, k is the discrete frequency, w represents the discrete window function [101].

It is essential to fulfill the requirement of the stationarity of the signal within a segment. Besides, it is important to keep in mind that the smaller the segment the higher the time resolution and the lower the frequency resolution. So a balance has to be made so that the size of the epochs is not to big to define the position of the transient event and that is not too small to compromise the spectrum of frequencies. Therefore the size of each epoch was set to one second containing 256 points each with an overlap of 50%.

The window w used to apply the spectrogram was a Kaiser window [102] of 256 points corresponding to 1 second and with a β value of 18, as seen in the Figure 4.8. The β characterizes the width of the curve defining for a specific point the influence of its neighbours in the spectrogram.

From the Matlab spectrogram it is possible to retrieve the PSD centers of energy and the correspondent time and frequency. The frequencies were only accepted in the specific band from [8-30Hz] every PSD point with a center of energy with an

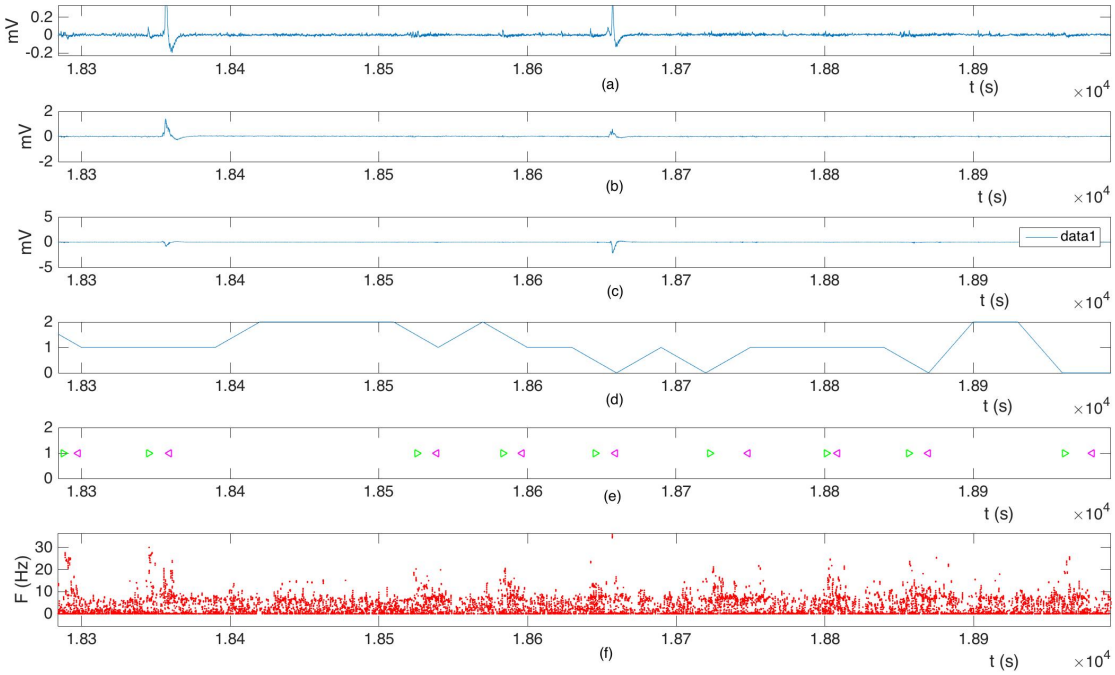


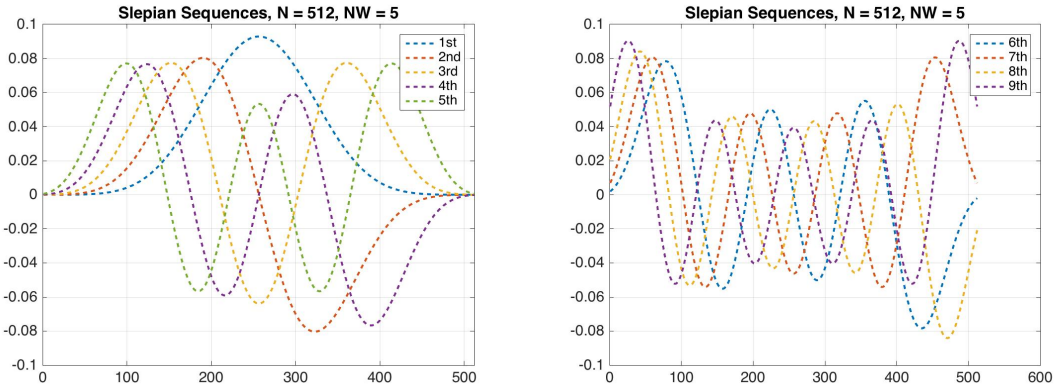
Figure 4.9: The three channels of the MESA database EEG Fz/Cz (a), Cz/Oz (b) and C4/M1 (c). The expert classification of the sleep stage, the hypnogram (d), the arousal visual scoring with start of arousal (green arrow) and end (red arrow) marked (e) and the result of the spectrogram with the respective time and frequency of the centers of energy (f).

higher or lower frequency were discarded. Then the features used for the detection were the respective frequency f_{CE} and time t_{CE} of a center of energy.

4.2.2.2 MultiTaper Technique

As aforementioned the spectral analysis is the quantitive approach for describing the biomedical signal at different frequencies in terms of the basic sinusoids that compose it. The signal power is a function of time and frequency. A new method to perform the spectral analysis is presented in this section, and it is called Multitaper. The Multitaper method was developed by David Thomson in the early 80's [103]. This technique has proven to be more robust and precise statiscally wise when compared the use of a single taper to estimate the spectrum [104, 105]. A novelty aspect of this thesis is that only in the recent years Multitaper has been applied to EEG signals.

To execute the Multitaper method it is necessary to follow some steps. Initially, computes a set of window functions. These windows are part of a class of functions named discrete prolate sheroidal sequence (DPSS) [106], also called Slepian



(a) **(b)**

Figure 4.10: The different slepian windows used in the multitaper method: (a)the first five discrete prolate spheroidal sequences of length 512 elements corresponding to 2 seconds; (b) from the sixth to the ninth slepian sequences.

sequences. The advantages of these tapers are diverse, the most important is the potential to reduce bias. Since they are all mathematically orthogonal functions, the spectral estimate retrieved from each taper is uncorrelated with others. The number of tapers is calculated by the following equations:

$$L = 2 \times TW - 1 \tag{4.5}$$

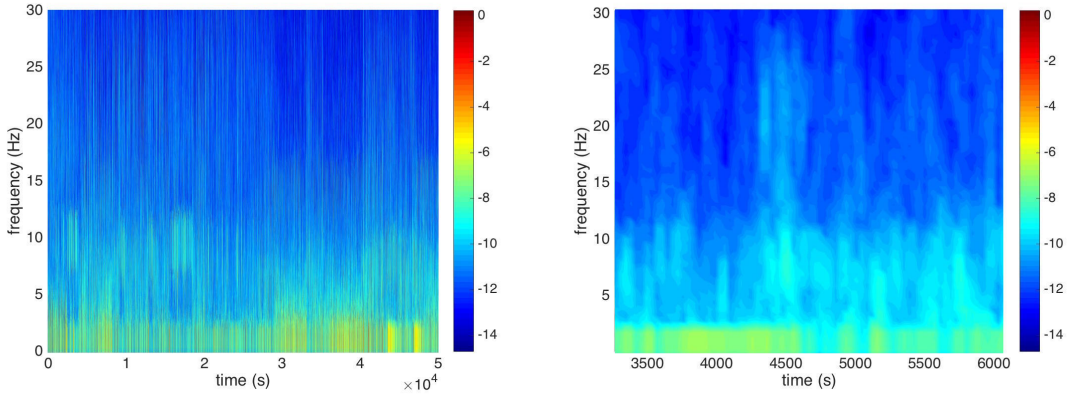
$$TW = N \times \Delta f/2 \tag{4.6}$$

where L is the number of tapers, TW is the time half-bandwidth, N the size of the window and Δf the frequency resolution [107].

Due to the fact that the use of the Multitaper to detect a microevent that occurs on a scale of a few seconds the window size was $N = 2s$ overlapped by 1 second , the frequency resolution was to be compromised to maintain low-variance. A frequency resolution $\Delta f = 5Hz$ was selected leading to, a number of tapes of 9 and their shapes can be seen in the Figure 4.10.

After, for each taper it is calculated the single-taper spectrum. The results can be averaged together as independent trials reducing the variance of the spectral analysis due to the fact that the tapers are uncorrelated. The final spectral analysis can be seen in the Figure 4.11.

The feature used for the detection is the sum of the power of frequency bands of



(a) **(b)**

Figure 4.11: The color map generated from the spectral analysis using Multitaper technique, representing the power in function of frequency and time. (a) the over all look of the signal and (b) a zoom of a short period of time. The first image tends to be segmented because the main point is to study transient events of small duration and not to classify long lasting events.

interest. Considering the band of frequencies [8 30Hz] the result of the sum of the powers can be seen in the Figure 4.12.

4.2.2.3 Rule-based Classifier

With the information retrieved from both detectors it is necessary to establish some rules to define in which conditions the arousals are detected. Besides, it is of most importance to follow the 11 scoring rules defined by the ASDA task force [18]:

1. The subject must be asleep for at least ten seconds before an arousal can be scored.
2. Two arousals have to be separated for at least 10 seconds of continuous sleep.
3. The duration of the candidate arousal has to be bigger than 3 seconds.
4. In NREM sleep arousals can be scored without an increase of the EMG amplitude.
5. In REM sleep arousals can only be score with an increase of the EMG amplitude.
6. A change of the EMG activity cannot be scored as an arousal.
7. Other transient events like K-complex and artifacts or even Delta waves cannot be classified as an arousal, except for some exceptions.

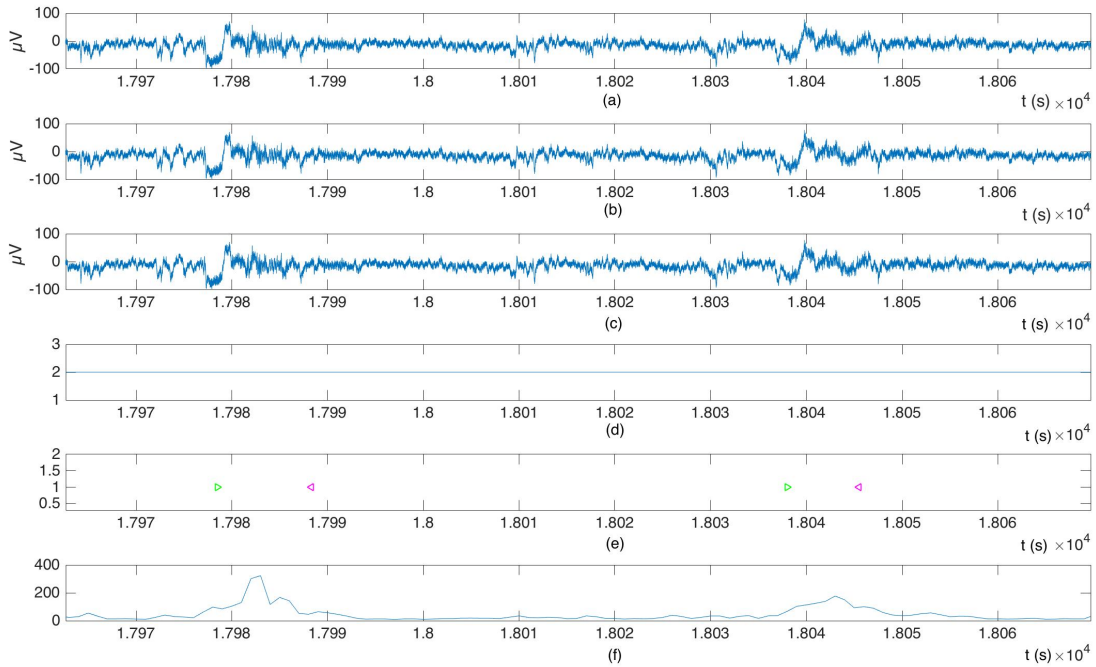


Figure 4.12: The three channels of the MESA database EEG Fz/Cz (a), Cz/Oz (b) and C4/M1 (c). The expert classification of the sleep stage, the hypnogram (d) the arousal visual scoring with start of arousal (green arrow) and end of arousal (red arrow) (e) and the result of the sum of the powers of the bands of frequencies [8 30Hz] as a function of time using Multitaper spectral analysis (f) .

8. If an arousal is contiguous with a pen blocking artifact they must be classified as an arousal.

9. If a non concurrent but contiguous EEG and EMG changes are in total bigger than 3 seconds and individually less than that, an arousal cannot be scored.

10. A three second long segment of sleep of alpha activity can only be scored as an arousal if it is followed by 10 segments free of alpha activity.

11. Transitions of sleep are not enough to score an arousal unless they comply with all the previous rules.

In this thesis the classification of the arousal was made with only the information provenient from the EEG thefore from the rules above only some can influence the arousal classification.

The arousal classifier started by imposing a minimum threshold, this value was empirically tested and differed with the database and with the method. All the segments where the hypnogram was zero, meaning that, the subject was awake were excluded following the ASDA rule 1. The first valid point above the threshold

was designated a *possible arousal start (PAS)* once a value below the threshold was detected the point contiguous before that was designated as a *possible arousal end (PAE)*. If the $\Delta t = PAE - PAS$ was bigger than three seconds following rule number 3, the arousal was validated. Then the PAS is saved as *arousal start (AS)* and PAE as *arousal end (AE)*. In order to define the next PAS it was imposed that the time within end of the previous scored arousal and the possible arousal start was bigger that 10 seconds, $PAS_n - AE_{n-1} > 10s$.

4.2.3 K-complex Detection

Matched filters were defined by George Turin in 1960 [60] and over the years many sleep researchers have benefited from this method ever since [4, 5, 64] for the detection of K-complex. It is of most importance to take into account the most defining characteristic of the K-complex, is its shape, so this thesis focus the detection of this transient event using Matched Filter technique.

4.2.3.1 Matched Filters

The Matched filter is a filter that compares a given shape or waveform to a signal, in this study we are comparing the template of a K-complex with the EEG signal. The Matched filter impulse response $h(\tau)$, is given by:

$$h(\tau) = ks(\Delta - \tau) \tag{4.7}$$

where $s(t)$ is the function of the waveform to detect, and k and Δ are arbitrary constants [60].

This filter is also called a "conjugate " filter because its transfer function is mathematically the complex conjugate of the spectrum of the signal to which it is matched. The challenge in using this filter for the detection of K-complex is the definition of the waveform or shape used for the detection. The template is essential to detect the transient events, it has started by having the perfect mathematical shape of the waveform [4], evolving to some more complex waveforms and even the combination of multiple shapes [5], as seen in figure 3.1.

The template used in this thesis was the result of the average shape of EEG signals classified by experts as a K-complex. First, the segments where the transient event was marked were selected. The position of the wave in the segment is not always

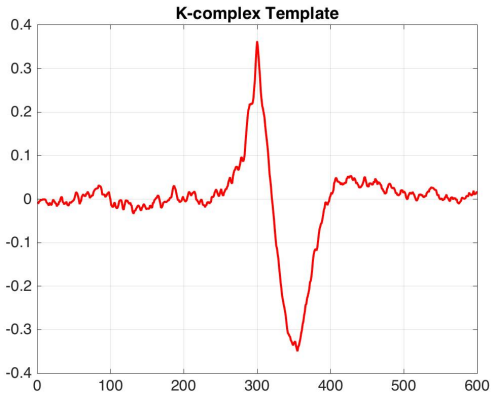


Figure 4.13: The template used for the matched filter method in the detection of K-complex in EEG signal.

the same, so a simple mathematical average would be a mistake and would ruin the template. So all the segments were centered based on their maximum point the peak and after averaged. The result is the template seen in the figure 4.13.

Finally, it is performed a convolution of the EEG signal with the template which retrieves a template in which the elements are the area below the points of the template sliding through the signal. The result can be seen in the Figure 4.14.

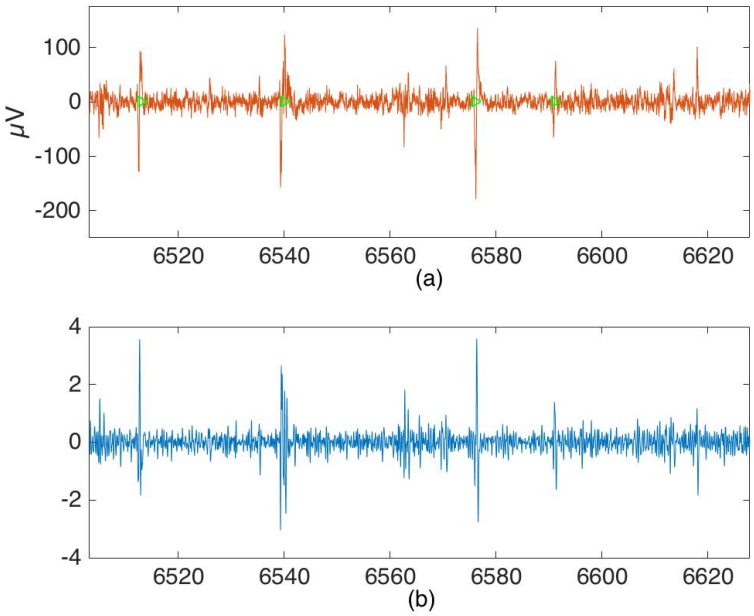


Figure 4.14: The result of the matched filter (b), applied to the independent EEG signal (a) to detected the visual scorings of the K-complexes (green arrows).

4.2.3.2 Rule-based Classifier

Unlike arousals, K-complex do not have many rules to define their validity, the shape of the signal is easily identified visually discarding the need to impose complex rules. Therefore the complexity of the classifier decreases. In this project, the classifier is simply a threshold value that is empirically tested to find the most appropriate one.

4.2.4 Performance Evaluation

To evaluate the performance of the detector it is necessary to introduce some statistical concepts that enable the classifier to be compared with the real classification. Considering the label with only two classes to be or not to be a specific transient event, and taking into consideration the result of the classifier, the combination of possible results is:

- *TP true positive* - when in the presence of a transient event the classifier detects it.
- *FP false positive* - when in the absence of a transient event the classifier detects one.
- *TN true negative* - when in the absence of a transient event there is no detection by the classifier.
- *FN false negative* - when in the presence of a transient event there is no detection by the classifier.

With the four possible outcomes, a two-by-two confusion matrix can be constructed representing the results of the classification. The confusion matrix is the first step to that evaluate the performance, and it can be seen in the table 4.3.

Table 4.3: Confusion Matrix of a Binomial problem. The numbers along the diagonal correspond to the correct decisions made by the classifier.

		Label	
		P	**N**
Results	**P**	TP	FP
	N	FN	TN

To evaluate the performance of the classifier four measures are used. Sensitivity, (SS) which is also called true positive rate, is the probability of truly detecting an event, the mathematical expression is the equation below:

$$SS = \frac{TP}{TP + FN} \tag{4.8}$$

Specificity (SP), is the percentage of non-detections that are correct, and the mathematical equation is the one below.

$$SP = \frac{TN}{TN + FP} \tag{4.9}$$

Accuracy (ACC), the measure of statistical bias, has the mathematical equation of:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.10}$$

False positive rate is the probability of s the probability of falsely detecting an event that is not truly present. Its equation is the one below.

$$FPR = \frac{FP}{TN + FP} = 1 - SP \tag{4.11}$$

To study the influence of the variance of the threshold within the classifier we introduced the ROC curve. A Receiver Operating Characteristic is a two-dimensional plot in which the SS is on the Y axis and the FPR is plotted in the X axis [108]. It allows to see the behaviour the classifier with the shape of the curve. The diagonal line with the equation $y(x) = x$ is the graphic representation of randomly guessing the presence of a transient event. A good classifier that is retrieving positive information from the EEG and is able to strategically detect events is situated in the upper region far from the random line.

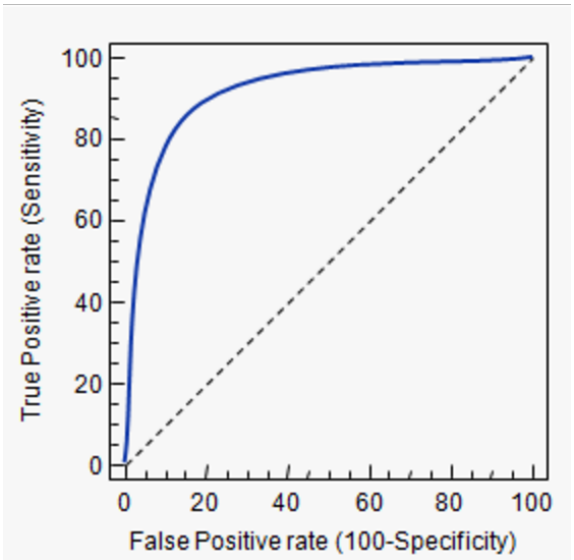


Figure 4.15: The plot of a ROC curve and the bisector line of the first quadrant.

To reduce this complex curve to a simple easily comparable number it is calculated the area under the ROC curve [109], or AUC. The AUC possible values are between [0 1] because AUC is a portion of the unit square. The random guess line now corresponds to a AUC value of 0.5 and no value bellow that is reasonable. Other characteristic of the AUC is that it is "equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance" [109]. In conclusion, the higher the AUC the better the classifier.

Chapter 5

Results

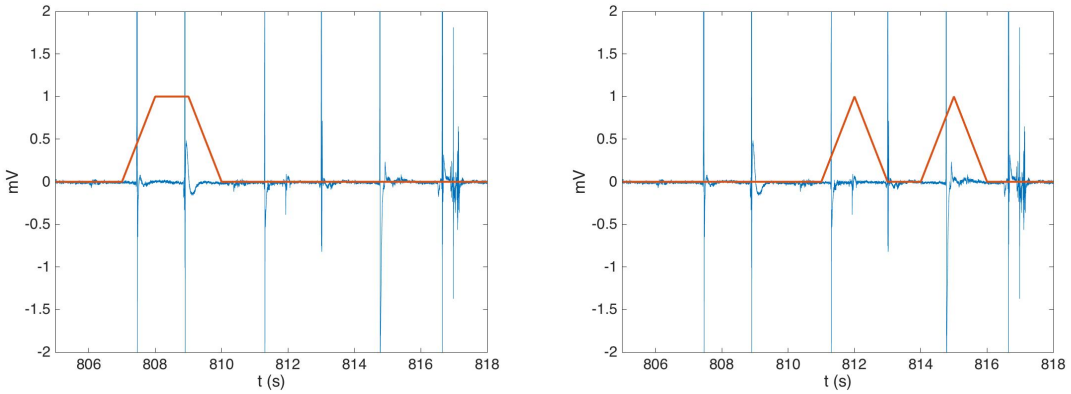
The goal of this project was to achieve a robust classifier of transient sleep events. In order to fulfil this goal three classifiers of transient events were developed, tested and compared. Two classifiers were used for the detection of arousals and one for the detection of K-complexes. In this chapter, the results achieved with the classification process are reported. In Section 5.1 it is reported the influence of the artifact detectors used in this project.

5.1 Artifact Detection

Aiming to improve the robustness of the classifiers, two artifact detectors were computed. Artifacts are any interference that is not provenient from the brain signals may be caused by muscular activity, movement of the eyes , or poor electrode content and in any case they need to be removed from the signal. Their detection is essential owing to the fact that an artifact may also produce an increase of the frequency and in the classification process this can be mistaken as an arousal. Beyond that, their

shape may resemble a K-complex but with bigger amplitude. There was no previous classifications of the artifacts so the classifiers were produced with a conservative approach to ensure that if an artifact is detected is actually a true artifact, reducing the probability of false detections and exclusion of good signal.

We report an example of use of the EEG spectral power for artifact detection in Figure 5.1, with two samples of a detection of an increase of slow waves, in Figure 5.1b and of the fast waves, in Figure 5.1a.



(a) **(b)**

Figure 5.1: The artifact classification from the Spectral detector by raise of the power of fast waves (a) and slow waves (b) for subject 1 from the MESA database.

An example of the use of the MSE artifact detector is reported in Figure 5.2. This technique allows to detect the artifact produced by the movements of the subject.

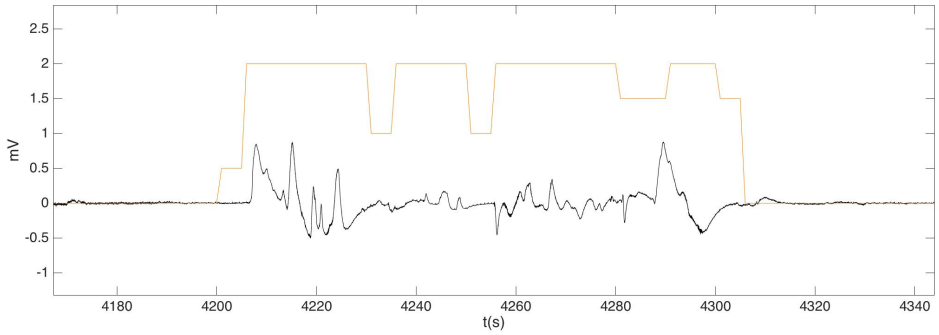


Figure 5.2: The artifact caused by movement of the subject classification from the MSE detector in a lighter colour.

Table 5.1: The total time classified as artifact by MSE and Spectral artifact detector for each subject of the MrOS database.

Subject	*MSE detected time (s)*	*Spectral detected time (s)*	*MSE ∩ Spectral(s)*	*Total (s)*
1	730	2220	315	2635
2	415	990	220	1185
3	1220	4380	550	5050
4	170	2729	375	2524
5	2299	4589	859	6029
6	495	2069	320	2244
7	4019	2579	1229	5369
8	1055	1830	470	2415
9	7190	4500	1290	10400
10	1579	4769	584	5764
11	640	2580	160	3060
12	1095	3870	570	4395
13	1379	2460	285	3554
14	1450	1080	275	2255
15	5414	3870	1260	8024
16	835	1980	320	2495
17	965	3270	215	4020
18	2455	3690	380	5765
19	1075	2520	325	3270
20	490	1380	145	1725

The artifact detections reduced the size of the signal where a classification could be done. After it the signal was not removed but with the ruled classifier it was imposed that no arousals can be found during these periods of time. The sum of the length of the artifacts present in each subject of the database MrOS and MESA are reported in Table 5.1 and Table 5.2, respectively.

The process of visually scoring an EEG signal is slow and time consuming therefore there is a possibility of wrong scores. Furthermore, sometimes arousals happen in concert with movements leading to a certain overlap in scoring. Considering this, the visual classifications were reviewed and the ones coincident with artifacts previously scored were considered invalid.

The total number of valid arousals for each subject changed and in the appendices is reported the valid visually scored arousals for both databases in the Table C.1. After this step the label used to inspect the quality of algorithm, is the validated visual score.

Table 5.2: The total time classified as artifact by MSE and Spectral artifact detector for each subject of the MESA database.

Subject	*MSE detected time (s)*	*Spectral detected time (s)*	*MSE ∩ Spectral(s)*	*Total (s)*
1	2595	4230	1370	5455
2	845	1710	385	2170
3	1460	2250	840	2870
4	805	2220	485	2540
5	1675	2640	965	3350
6	17190	2970	1975	18185
7	920	1770	550	2140
8	265	750	170	845
9	1695	4650	885	5460
10	685	2519	355	2849
11	1830	2250	775	3305
12	610	2310	415	2505
13	1855	2130	705	3280
14	1785	1020	185	2620
15	1280	2160	650	2790
16	1535	1769	644	2660
17	3095	1710	725	4080
18	5345	3180	1255	7270
19	2229	1920	460	3689
20	3700	3690	1085	6305

5.2 Arousal Detection

In this section the results from the detection of arousals are reported. Both techniques presented in the methodologic section are applied.

5.2.1 Spectrogram

First, the thresholds were tested using binary classification of the presence or absence of the arousal. With the variation of the threshold we generated the ROC curves to evaluate the classifier and to test the ideal threshold value. For the spectrogram the thresholds tested ranged from 0 to 20 for the MESA dataset and -59 to -40 for the MrOS dataset.

We report in Table 5.3 the values of average and standard deviation of the AUC, sensitivity, specificity, accuracy and the threshold value of the classification using

the spectrogram method on the whole dataset for both MrOS and MESA. Subset I and subset II differ on the number of subjects, subset II does not include the three subjects that are outliers(see Appendix) . The values of sensitivity, specificity and accuracy correspond to the threshold that have the minimum distance to the perfect recognition point (0,1) of the ROC graph.

The outliers were excluded based on corruption of the signal more details are reported in appendices in the Appendix A.

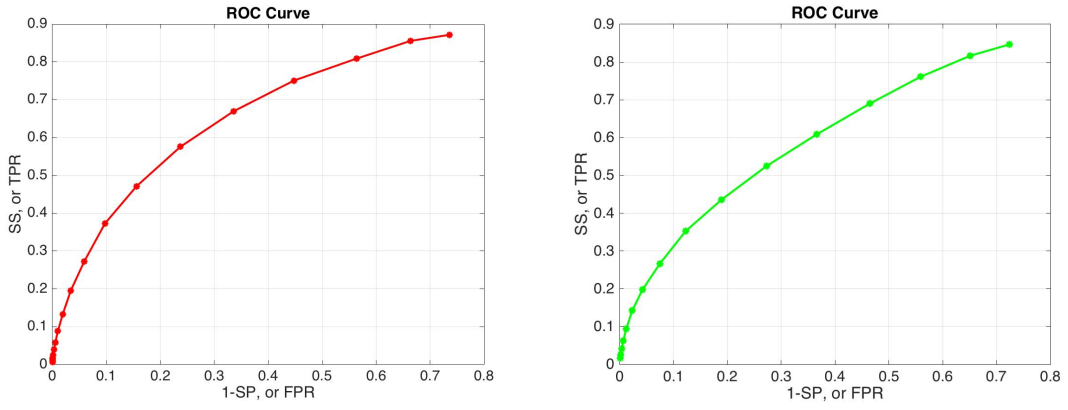
The results for every subject of the MrOS dataset are reported in the Table C.2 of the appendices, and the statistical variables for the MESA dataset subjects are reported in Table C.3 .

Table 5.3: Detection statistics obtained on both datasets for the two different subsets when using the spectrogram technique.

		AUC	*SP*	*SS*	*ACC*	*Threshold*
MrOS	*I*	0.759 ± 0.108	0.700 ± 0.098	0.721 ± 0.104	0.703 ± 0.095	6.7 ± 1.9
	II	0.804 ± 0.057	0.731 ± 0.062	0.751 ± 0.068	0.733 ± 0.060	7.1 ± 1.6
MESA	*I*	0.743 ± 0.071	0.702 ± 0.064	0.676 ± 0.083	0.686 ± 0.068	-51.8 ± 3.6
	II	0.798 ± 0.037	0.701 ± 0.053	0.694 ± 0.047	0.700 ± 0.047	-52.1 ± 3.8

The MrOS dataset has a higher average AUC, specificity, sensitivity and accuracy, although the standard deviation is higher too. This suggests the subjects present a higher variability when compared to the subjects of the MESA dataset.

We report an average ROC curve generated by the average of each subject's data SS and FPR for all the thresholds in Figure 5.3.



(a) **(b)**

Figure 5.3: The average ROC curve of the MrOS (a) and MESA (b) for the Spectrogram detector.

Both of the ROC curves are fairly distant from the bisector of the first quadrant. It can be noticed also that both the curves get close to the perfect recognition point (0,1), but the closest is the one relative to MrOS dataset.

In the appendices section are reported all the ROC curves for subjects of the MrOS dataset in Figure B.1 and in Figure B.2 the ones correspondent to MESA dataset.

The second part of our analysis consisted of imposing the AASM rules for arousal definition. In this case we evaluated the performance of our classifier by comparing the scored arousals as a whole and not one-second size epochs, thus the number of FN cannot be calculated. In Table 5.4 , it is reported the ratio of correct detections(sensitivity), and the number of falsely detected arousals. The ratio of correct detections is the quotient of the number of true detections over the number of valid visual scores.

Table 5.4: Detection statistics obtained on both datasets for the two different subsets and two different threshold values when using spectrogram technique and applying the decision rules.

		Average Threshold		*Ideal Threshold*	
		SS	*# False detections*	*SS*	*# False detections*
MrOS	*I*	0.571 ±0.235	195 ± 140	0.780 ± 0.106	143 ± 72
	II	0.604 ± 0.205	201 ± 134	0.808 ±0.068	147 ± 70
MESA	*I*	0.539 ± 0.312	311 ± 312	0.732 ± 0.171	301 ± 110
	II	0.581 ± 0.306	352 ± 320	0.740 ± 0.046	284 ± 108

The ideal threshold for each subject is the value that minimizes the proximity to the maximum recognition point before calculated. The average threshold is calculated from the average of all the ideal thresholds of each subject.

The ratio of correct detections is significantly higher when using the ideal threshold characteristic of each subject. The MrOS dataset seems to have higher average results when compared to the MESA dataset.

5.2.2 Multitaper Technique

Similarly to what was done before, initially the thresholds were tested using binary classification of the presence or absence of the arousal using as an only rule the threshold limit.

In Table 5.5 the values of average and standard deviation of the AUC, sensitivity, specificity, accuracy and the optimal threshold value of the classification using the

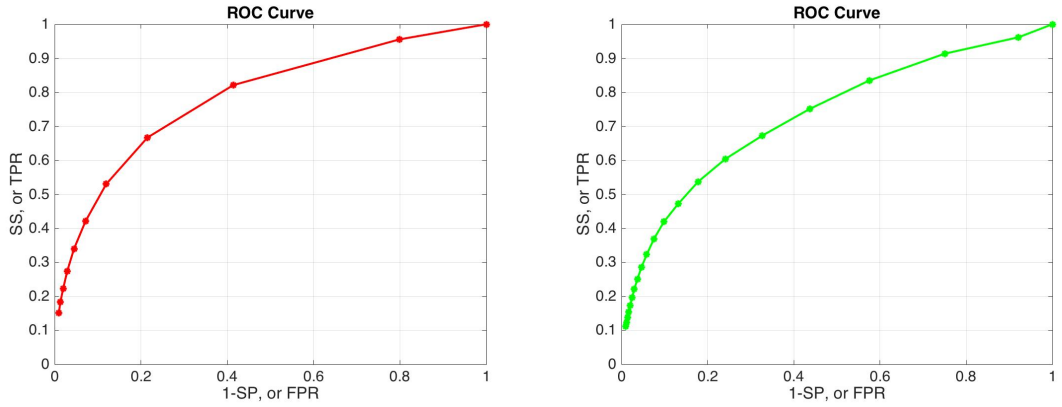
multitaper method on the whole dataset for both MrOS and MESA, are reported. Subset I and Subset II correspond to the same tests used in the spectrogram method removing the same outliers. The values of sensitivity, specificity and accuracy were calculated using the same method as before.

Table 5.5: Detection statistics obtained on both datasets for the two different subsets when using the multitaper technique.

		AUC	*SP*	*SS*	*ACC*	*Threshold*
MrOS	*I*	0.822 ± 0.089	0.777 ± 0.081	0.761 ± 0.079	0.770 ± 0.078	29.4 ± 12.4
	II	0.853 ± 0.050	0.799 ± 0.054	0.789 ± 0.072	0.795 ± 0.052	32.0 ± 12.2
MESA	*I*	0.792 ± 0.112	0.766 ± 0.059	0.729 ± 0.013	0.746 ± 0.079	53.1 ± 29.7
	II	0.833 ± 0.040	0.780 ± 0.042	0.775 ± 0.311	0.772 ± 0.041	59.5 ± 27.3

The MrOS dataset has a higher average AUC, specificity, sensitivity and accuracy, with a lower standard deviation , when compared with the other dataset.

In Figure 5.4 we report an average ROC curve generated by the average of each subject's data SS and FPR for all the thresholds.



(a) **(b)**

Figure 5.4: The average ROC curve of the MrOS (a) and MESA (b) for the Multitaper detector.

Both of the ROC curves are fairly distant from the bisector of the first quadrant. It can be noticed also that both the curves get close to the perfect recognition point (0,1), but the closest is the one relative to MrOS dataset.

In the appendices section are reported all the ROC curves for subjects of the MrOS dataset in Figure B.1 and in Figure B.2 the ones correspondent to MESA dataset.

In Table 5.6, it is reported the ratio of correct detections, and the number of false arousals. The ratio of correct detections is the quotient of the number of true detec-

tions over the number of valid visual scored arousals. All the above are calculated after imposing the AASM arousal detection rules.

For the first part of testing the classifiers of arousals generate the ROC curve for each subject, the curves are in the appendices in the Figure B.1 and Figure B.2.

Table 5.6: Detection statistics obtained on both datasets for the two different tests and two different threshold values when using multitaper technique applying the decision rules.

		Average Threshold		*Ideal Threshold*	
		SS	*# False detections*	*SS*	*# False detections*
MrOS	*I*	0.600 ± 0.236	182 ± 143	0.750 ± 0.152	155 ± 76
	II	0.664 ± 0.164	189 ± 131	0.803 ± 0.089	163 ± 75
MESA	*I*	0.595 ± 0.34	301 ± 331	0.793 ± 0.153	277 ± 105
	II	0.674 ± 0.320	348 ± 342	0.833 ± 0.125	256 ± 107

The sensitivity is significantly higher when using the ideal threshold characteristic of each subject. The MESA dataset seems to have higher average results when compared to the MrOS dataset.

The results achieved when using the Multitaper are now compared with the ones from Spectrogram. For MrOS in subset two it was achieved a SS of 0.803 while in Spectrogram the SS value was of 0.808, so slightly higher in the Spectrogram. For the MESA database in the same conditions spectrogram achieved an SS of 0.740 and Multitaper a SS of 0.833, so considerably higher for the Multitaper.

5.3 K-complex detection

In the MrOS and MESA datasets used there was no visual scoring of K-complexes. One independent signal was visually classified in three different segments and those were used to train and test the detector in this first phase of this analysis.

We report the ROC curve of the K-complex detector applied to this EEG signal, in the Figure 5.5, when testing the threshold values between 0 and 4.

This ROC curve as an AUC of 0.814 and the value closer to the perfect recognition point (0,1) has a sensitivity of 0.730 and specificity of 0.819.

Due to having no visual scores of K-complexes on the datasets under analysis we performed a qualitative assessment of the performance of our detector. In addition, we report the distribution of scored K complexes by sleep stage.

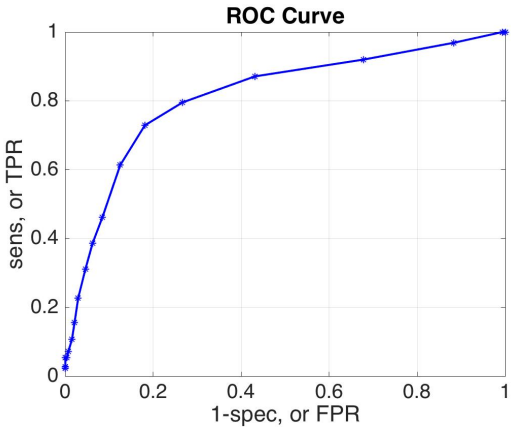


Figure 5.5: The average ROC curve of the K-complex detection.

In the Figure 5.6 is reported an example of two detections using this macthed filter technique in the MESA dataset.

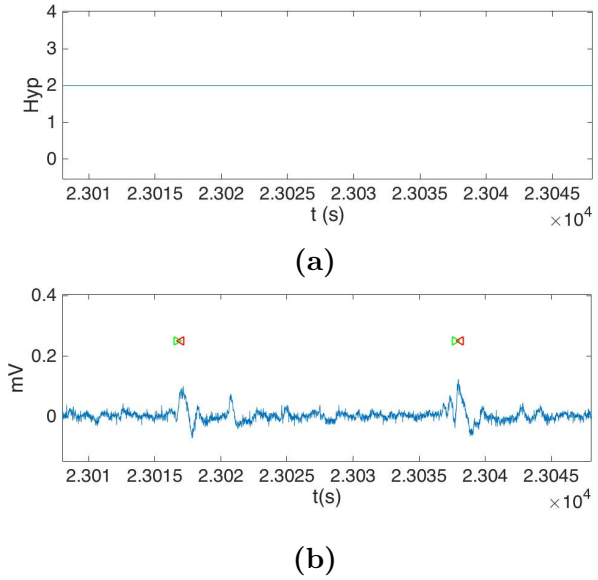
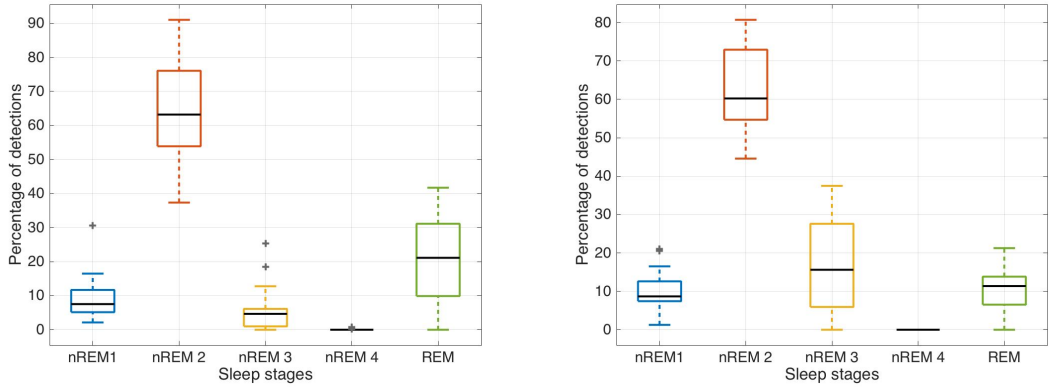


Figure 5.6: An example of a K-complex detection in the EEG signal (b), with the respective hypogram value (a), for the subject 1 from the MESA dataset.

In the Figure 5.7 is reported the boxplot of the results of the K-complex detection for both dataset MrOS and MESA, in 5.7a and 5.7b respectively, with an empirically threshold equal to 1.9 .



(a) **(b)**

Figure 5.7: The boxplots of the ratio of K-complexes for the five sleep stages for the datasets from MrOS (a) and MESA (b).

The percentage of K-complex detected is higher for both datasets in the Sleep nREM 2. In the MESA boxplot the second highest average is nREM 3 which in MrOS is REM.

Chapter 6

Discussion

This study provides an efficient way to detect two types of transient sleep events, K-complexes and arousals. The sample used for this study was composed of 40 subjects from two different databases, with more than 8 hours of sleep each, and a EEG sample frequency of 256 Hz. This dataset is considered large enough to allow the results to have significant importance. The diversity of the datasets and the use of a different channel for each database allows to test the robustness of the classifier.

The use of the artifact detectors allowed the classifiers to avoid false detections of arousals, due to the raise of frequency present in some of the artifacts, and K-complex, due to the similarity of the shape shared by the two sleep structures. The use of artifact detectors is common in EEG analysis [57, 50] but most studies use only filters to remove the powerline interference [58]. Over the literature reviewed in this thesis, to my knowledge, only Agarwal et al. used a ECG interference detector, [49]. This is the first combination of three classifiers and it fulfils the needs of removing noise from EEG signals provinient from different sources.

The MSE technique and the use of the spectral analysis to detect artifacts comple-

ment each other, their theoretical principle is different and they can detect differences of entropy (regularity) of the signal and changes in the frequency spectrum, respectively. In the Table 5.1 and Table 5.2 the total duration of the artifacts detected by both classifiers is reported, for both datasets, MrOS and MESA respectively. They show that although there are some artifacts classified by both detectors, there is always a certain period of time that is only detected by one of the classifiers. Mathematically it can be seen because the total time of detection that results from the intersection of both classifiers is always inferior to the total time of detection of any single classifier.

Another conclusion that can be drawn from the detection of artifacts is the validity of the signals. Three signals from each database were excluded due to being corrupted in some way. One of the signals was provenient from the subject 6 from MESA database. As it can be seen, in Table 5.2 it has a total time of artifacts with a bigger order of magnitude when compared to the rest. Another potential use of this artifact detection procedure would be to check the validity of the labelled arousals.

Unlike other sleep-related events such as spindles, arousals are not as well defined. The processes of visually scoring an arousal is very subjective besides it is a very long and time-consuming process that can easily lead to mistakes. So the visually scored arousals were validated to make sure they were not present neither in a awake stage neither in a segment of a signal classified as an artifact, the results can be seen in the Table C.1. This step is essential for two reasons: first, because the label may actually be a mistake; second, because our classifiers would no be able to detect this arousal due to the imposed rules upon the classification process.

For the Spectrogram is clear in the Table 5.3 that statistical values improve after excluding the subjects that were considered invalid. The AUC for MrOS increases from 0.759 to 0.804 and in MESA from 0.743 to 0.798. It is important to notice that the standard deviation is also reduced. The results from all subjects can be seen in the Table C.2 and Table C.3. The results for the MrOS dataset are all better when compare with the ones achieved for the MESA dataset, this may be due to the expert classification, use of a different channel, or to the quality of the acquisition of the signal.

The average ROC curve in Figure 5.3 reporting the average value for each threshold is also better (higher AUC) for the MrOS dataset.

For the Second part, reported in the Table 5.4, the use of an average threshold for all subjects resulted in lower performance when compared to the first part of

this study when it was not imposed AASM rules. When using the ideal thresholds characteristic of each subject the results are considerably higher, achieving for MrOS a SS of 0.808 and for MESA 0.740.

Once again after excluding the invalid subjects subset II the results are better for the same reasons as before. For the MrOS dataset the subset II has a SS of 0.808 while the other subset has a SS of 0.780, for MESA the results were 0.732 and 0.740 respectively. When using the ideal threshold the results are better compared to the average threshold, and to the first part of the study. The reason for this is because when using one second epochs the visual classification is not precise and our arousal is never 100% coincident so the extra seconds that the arousals are not coincident decrease the overall score. The number of false detections is very high, this is a consequence of using only EEG. The AASM rule number four states that an arousal can only be visual classified with an increase of the EMG amplitude, so if a raise of frequency of the EEG signal is detected by the classifier may not be a visually scored as an arousal for the lack of the change of the EMG signal.

The multitaper method was also tested in the same way. The first statistical results are represented in the Table 5.5. The statistical results for MrOS were all higher in subset II. As an example,the average AUC of subset I is 0.822 and the AUC for subset II is 0.853. This dataset achieves better result than the other dataset, MESA. The AUC of subset II in of MrOS is 0.853 while for MESA is 0.833. The average ROC curve in Figure 5.4 reporting the average value for each threshold is also better (higher AUC) for the MrOS dataset.

In the second part of the study (Table 5.6) comparing the ratio of correct detections with the sensitivity of the first part of the study the results improve intra-dataset. Unlike what happens with the other classifier the MESA achieves better results on the ratio of correct detections but with a bigger standard deviation. In summary the Multitaper technique has higher values in almost all statistical variables compared to the traditional spectral analysis. The spectrogram produces high variance across all the frequencies that are calculated representing there is a compromise between time and frequency resolution, while the multitaper method has a better definition of short time events because the number of tapers can compensate the small number of data used to calculate the spectrum.

The models which use only EEG signal for the arousal classification and the study of Agarwal et al. that combines the information of the ECG presented in the literature are considered and compared with the Spectrogram and Multitaper method used in this thesis in Table 6.1.

Table 6.1: The statistical results of the presented literature in comparison with the results achieved in this thesis.

	Method	*AUC*	*SS*	*SP*	*ACC*
Agarwal et al. [49]	Adaptive segmentation + Decision rules	-	0.822	0.724	
Cho et al. [50]	TFA + SVM	-	0.753	0.938	-
Shmiel et al. [51]	Critical Points + MetaRules + Decision rules	-	0.494	-	0.825
Liang et. al [56]	Curious Extreme Learning Machine	0.79	-	-	0.85
Thesis	Spectrogram + Decision rules	-	0.808	-	-
	Multitaper + Decision rules	-	0.833	-	-

When comparing the results achieved in this thesis with the results achieved by Agarwal et al.[49], the other author has a higher SS of 0.822 when compared to the spectrogram, SS of 0.808 , but lower when compared with the multitaper, SS of 0.833 . However, their study used only two subjects for all the testing and were visually scored by three experts each. While, in this thesis each subject from the dataset was labelled by only one subject. The high inter-expert variability influences our study [21] and also intra-expert variability seems to be high when performed over time , and this also happens in our study too. All the other experts that use the only EEG channels [50, 51] have worse results when it comes to the SS achieved. Cho et al. achieved a SS of 0.753 and Shimel et al. achieved a SS of 0.494. Although better results were achieved by using EMG combined with the EEG, that crosses the goal of this thesis.

For the detection of K-complexes is performed a matched filter technique. In the first part of our analysis we used a signal with annotated K-complexes in this case it is achieved a ROC curve with a AUC of 0.8127. The other authors that performed K-complex detection did not calculate AUC values so the SS and the SP are the values that are going to be discussed in this section. We are going to compare with the previous studies that report a maximum SS and SP. In the Table 6.2 the results achieved in this classifier are compared with this thesis.

Comparing the results with the ones from the Literature the Woertz study in the first performed technique, TFA achieved a SS and SP of 0.81 and 0.93 both higher than the results achieved in this thesis, the second study with matched filters has a higher SS of 0.93 and a lower SP 0.68 when compared with this thesis. Woertz et

Table 6.2: The statistical results of the K-complex detection from the presented literature in comparison with the results achieved in this thesis.

	Method	*SS*	*SP*
Woertz et al. [61]	TFA	0.81	0.93
	Matched filters	0.93	0.68
Shete et al. [64]	ANN	0.9606	0.5262
Erdamar et al. [70]	TEO + Wavelets	0.85	0.93
Thesis	Matched filters	0.730	0.819

al. [61] study employs 10 min epochs classified by 3 experts. The results achieved in the Shete et al. study have a SS of 0.9606 and a SP of 0.5262 although the SS is higher than the one achieved in this study the SP is significantly lower than the achieved in this thesis. Erdamar et al. achieved a SS of 0.85 and a SP of 0.93: both results are better than the ones achieved in this thesis, this study uses the results from three expert visual scoring. The results from this thesis could be improved by instead of comparing the 1-second epoch classification to consider the K-complex as one single score and to check if the transient event has been detected or not. Moreover, it would be good to increase the number of signals used for the training of the detector and the number of experts that do the scoring of this transient event.

The K-complex detection was applied to the MESA and MrOS database and one example of a good classificaton is shown in the results section in Figure 5.6. The beginning of the K-complex (green arrow) is correctly scored although the end of the transient event (red arrow) is not in the right position, which also decreases the statistical results achieved. It would be relevant to adjust this value.

Finally, Figure 5.7 reports the distributions of the scored K-complexes by sleep stage. These boxplots report the percentage of detections achieved by using the matched filter detection in our database. As expected the percentage of detections is higher in the sleep stage nREM 2 (orange box), due to this transient event being an essential feature of this sleep stage. K-complex have a very similar shape to the vertex waves, present in the sleep stage nREM 3 so the second highest box would be expected to be in this sleep stage, but it only happens for the MESA database. The MrOS is a dataset encompassing elderly men, thus it has a smaller percentage of detections during this sleep stage. The deep sleep, nREM 3 is usually decreased with aging, so that influences the number of detections [13].

Chapter 7

Conclusions and Future Work

The main goal of this thesis was to develop robust classifiers of transient events on sleep EEG, in specific, arousals and K-complexes.

This thesis presents multiple classifiers for the transient events of sleep, arousals and K-complex. The novelty introduced in this study was the use of MSE and spectral analysis for the detection of artifacts and ECG artifact remover, which improves clearly our results and heightens the robustness of the classifier

Two different arousal classifiers were tested and it is possible to conclude that the multitaper method is the best for the detection of the arousals. This method has the ability to reduce the bias and the variance when compared to the Spectrogram. It achieved a SS of 0.833 and its performance is better than that reported in our studies literature, using only a information prevenient from the EEG signal. The high inter-expert variability influences our study [21] and also intra-expert variability seems to be high when performed over time, and this also happens in our study.

We can conclude that the methods explored in this thesis seem to have the potential for developing a robust classifier of EEG arousals. It must be noted that "ideal

threshold" for each individual is still the best way to achieve good classifications, while using average thresholds reduces the performance. In order to develop a classifier of new data, based on a single, universal threshold, in the future would be beneficial to normalize the signals by their standard deviation, or to generate an algorithm to calculate a threshold adapted for each subject.

The K-complex detector achieves good statistical results that are lower than some from the literature. The SS achieved is significant lower than the SP but both are good results considering the situations on which this detector was developed. The classifier seems to have a good robustness, because it was applied to both datasets and although there are no gold standard annotations, the results are acceptable by visual inspection. Furthermore the high density of K-complexes in stage nREM 2 compared to other stages represents a reassurance of the accuracy of the method.

Taking into consideration all the results achieved, we can conclude that our main goal was met. Future steps will involve the improvement and further testing of the K-complex classifier, as well as the integration of a spindle detector. In the near future, it is intended the development GUI for public use, that includes the classification of these three transient events. This way giving to any non expert the possibility of inspecting the sleep EEG.

Bibliography

[1] I. Bankman, V. Sigillito, R. Wise, and P. Smith, "Feature-based detection of the k-complex wave in the human electroencephalogram using neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 12, p. 1305–1310, 1992.

[2] G. W. Don and K. A. Waters, "Influence of sleep state on frequency of swallowing, apnea, and arousal in human infants," *Journal of Applied Physiology*, vol. 94, no. 6, pp. 2456–2464, 2003.

[3] S. Sanei and JonathonChambers, *EEG Signal Processing*. John Wiley & Sons, 2011.

[4] G. Bremer, J. R. Smith, and I. Karacan, "Automatic detection of the k-complex in sleep electroencephalograms," *IEEE Transactions on Biomedical Engineering*, vol. BME-17, no. 4, p. 314–323, 1970.

[5] N. Kerkeni, L. Bougrain, M. H. Bedoui, F. Alexandre, and M. Dogui, "Reconnaissance automatique des grapho-éléments temporels de l'électroencéphalogramme du sommeil," *5ème édition des Ateliers Traitement et Analyse de l'Information : Méthodes et Applications - TAIMA 2007*, May 2007.

[6] J. J. Vidal, "Toward direct brain-computer communication," *Annual Review of Biophysics and Bioengineering*, vol. 2, no. 1, pp. 157–180, 1973.

[7] S. L. Wendt, P. Welinder, H. B. Sorensen, P. E. Peppard, P. Jennum, P. Perona, E. Mignot, and S. C. Warby, "Inter-expert and intra-expert reliability in sleep spindle scoring," *Clinical Neurophysiology*, vol. 126, no. 8, p. 1548–1556, 2015.

[8] L. Parrino, P. Halasz, C. A. Tassinari, and M. G. Terzano, "Cap, epilepsy and motor events during sleep: the unifying role of arousal," *Sleep Medicine*

Reviews, vol. 10, no. 4, p. 267–285, 2006.

[9] A. Kales and A. Rechtschaffen, *A manual of standardized terminology, techniques, and scoring system for sleep stages of human subjects*. Brain Information Service/Brain Research Institute, University of California, 1968.

[10] C. Sleep and E. R. Center, *MrOS sleep PSG Procedure Manual*. Case Western Reserve University, 2004.

[11] C. Iber, *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*. American Academy of Sleep Medicine, 2007.

[12] K. Crowley, J. Trinder, Y. Kim, M. Carrigton, and I. M. Colrain, "The effects of normal aging on sleep spindle and k-complex production," *Clinical Neurophysiology*, vol. 113, no. 10, p. 1615–1622, 2002.

[13] M. Boselli, l. Parrino, A. Smerieri, and M. G. Terzano, "Effect of age on eeg arousals in normal sleep," *Sleep*, p. 361–367, Jan 1998.

[14] N. Wolkove, O. Elkholy, M. Baltzan, and M. Palayew, "Sleep and aging: 1. sleep disorders commonly found in older people," *Canadian Medical Association Journal*, vol. 176, no. 9, p. 1299–1304, 2007.

[15] K. Wulff, S. Gatti, J. G. Wettstein, and R. G. Foster, "Sleep and circadian rhythm disruption in psychiatric and neurodegenerative disease," *Nature Reviews Neuroscience*, vol. 11, no. 8, p. 589–599, 2010.

[16] A. Horvath, X. Montana, J.-P. Lanquart, P. Hubain, A. Szucs, P. Linkowski, and G. Loas, "Effects of state and trait anxiety on sleep structure: A polysomnographic study in 1083 subjects," *Psychiatry Research*, vol. 244, p. 279–283, 2016.

[17] B. J. Swihart, B. S. Caffo, C. M. Crainiceanu, and N. M. Punjabi, "Characterizing sleep structure using the hypnogram," *Journal of clinical Sleep Medicine*, vol. 4, no. 4, p. 349–355, 2008.

[18] ASDA, "Eeg arousals: Scoring rules and examples. a preliminary report from the sleep disorders atlas task force of the american sleep disorders association," *Sleep*, p. 174–184, 1992.

[19] M. G. Terzano, L. Parrino, A. Sherieri, R. Chervin, S. Chokroverty, C. Guilleminault, M. Hirshkowitz, M. Mahowald, H. Moldofsky, A. Rosa, and et al., "Atlas, rules, and recording techniques for the scoring of cyclic alternat-

ing pattern (cap) in human sleep," *Sleep Medicine*, vol. 2, no. 6, p. 537–553, 2001.

[20] M. J. Drinnan, A. Murray, C. J. Griffiths, and G. J. Gibson, "Interobserver variability in recognizing arousal in respiratory sleep disorders," *American Journal of Respiratory and Critical Care Medicine*, vol. 158, no. 2, p. 358–362, 1998.

[21] M. J. Drinnan, A. Murray, J. E. S. White, A. J. Smithson, C. J. Griffiths, and G. J. Gibson, "Automated recognition of eeg changes accompanying arousal in respiratory sleep disorders," *Sleep*, vol. 19, no. 4, p. 296–303, 1996.

[22] M. Terzano, D. Mancia, M. R. Salati, G. Costani, and A. Decembrino, "The cyclic alternating pattern as a physiologic component of normal nrem sleep," *Sleep*, Jan 1985.

[23] M. Roth, J. Shaw, and J. Green, "The form, voltage distribution and physiological significance of the k-complex," *Electroencephalography and Clinical Neurophysiology*, vol. 8, no. 3, p. 385–402, 1956.

[24] L. Johnson and W. E. Karpan, "Autonomic correlates of the spontaneous k-complex," *Psychophysiology*, vol. 4, no. 4, p. 444–452, 1968.

[25] P. Halasz, M. Terzano, L. Parrino, and R. Bodizs, "The nature of arousal in sleep," *Journal of Sleep Research*, vol. 13, no. 1, p. 1–23, 2004.

[26] F. Amzica and M. Steriade, "Cellular substrates and laminar profile of sleep k-complex," *Neuroscience*, vol. 82, no. 3, p. 671–686, 1997.

[27] T. Andrillon, Y. Nir, R. J. Staba, F. Ferrarelli, C. Cirelli, G. Tononi, and I. Fried, "Sleep spindles in humans: Insights from intracranial eeg and unit recordings," *Journal of Neuroscience*, vol. 31, p. 17821–17834, Jul 2011.

[28] M. G. Terzano and L. Parrino, "Origin and significance of the cyclic alternating pattern (cap)," *Sleep Medicine Reviews*, vol. 4, no. 1, p. 101–123, 2000.

[29] A. Ayoub, D. Aumann, A. Hörschelmann, A. Kouchekmanesch, P. Paul, J. Born, and L. Marshall, "Differential effects on fast and slow spindle activity, and the sleep slow oscillation in humans with carbamazepine and flunarizine to antagonize voltage-dependent na and ca2 channel activity," *Sleep*, Jan 2013.

[30] H. G. Wei, E. Riel, C. A. Czeisler, and D.-J. Dijk, "Attenuated amplitude of circadian and sleep-dependent modulation of electroencephalographic

sleep spindle characteristics in elderly human subjects," *Neuroscience Letters*, vol. 260, no. 1, p. 29–32, 1999.

[31] S. M. Fogel and C. T. Smith, "The function of the sleep spindle: A physiological index of intelligence and a mechanism for sleep-dependent memory consolidation," *Neuroscience n Biobehavioral Reviews*, vol. 35, no. 5, p. 1154–1165, 2011.

[32] B. Clemens and A. Ménes, "Sleep spindle asymmetry in epileptic patients," *Clinical Neurophysiology*, vol. 111, no. 12, p. 2155–2159, 2000.

[33] A. Castelnovo, A. D'Agostino, C. Casetta, S. Sarasso, and F. Ferrarelli, "Sleep spindle deficit in schizophrenia: Contextualization of recent findings," *Current Psychiatry Reports*, vol. 18, no. 8, 2016.

[34] M. H. Kryger, T. Roth, and W. C. Dement, *Principles and practice of sleep medicine*. Elsevier/Saunders, 2005.

[35] E. R. Sun, C. A. Chen, G. Ho, C. J. Earley, and R. P. Allen, "Iron and the restless legs syndrome," *Sleep*, vol. 21, no. 4, pp. 381–387, 1998.

[36] P. F. Diez, V. Mut, E. Laciar, and E. Avila, "A comparison of monopolar and bipolar eeg recordings for ssvep detection," *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2010.

[37] D. Kasper, A. Fauci, S. Hauser, D. Longo, L. Jameson, and J. Loscalzo, *Harrison's Principles of Internal Medicine*. McGraw Hill, 19 ed., 2015.

[38] G. Bodenstein and H. Praetorius, "Feature extraction from the electroencephalogram by adaptive segmentation," *Proceedings of the IEEE*, vol. 65, no. 5, p. 642–652, 1977.

[39] R. Agarwal and J. Gotman, "Adaptive segmentation of electroencephalographic data using a nonlinear energy operator," *ISCAS99. Proceedings of the 1999 IEEE International Symposium on Circuits and Systems VLSI*, 1999.

[40] S. Anisheh and H. Hassanpour, "Adaptive segmentation with optimal window length scheme using fractal dimension and wavelet transform," *system (stationarity)*, vol. 1, no. 4, p. 5, 2009.

[41] O.-D. Creutzfeldt, G. Bodenstein, and J. Barlow, "Computerized eeg pattern classification by adaptive segmentation and probability density function classification. clinical evaluation," *Electroencephalography and Clinical Neurophysiology*, vol. 60, no. 5, p. 373–393, 1985.

[42] M. Zamora, "The study of micro-arousals using neural network analysis of the eeg," *9th International Conference on Artificial Neural Networks: ICANN 99*, p. 625–630, 1999.

[43] B. Yegnanarayana, "Artificial neural networks for pattern recognition," *Sadhana*, vol. 19, no. 2, p. 189–238, 1994.

[44] F. D. Carli, L. Nobili, P. Gelcich, and F. Ferrillo, "A method for the automatic detection of arousals during sleep," *Sleep*, vol. 22, no. 5, p. 561–572, 1999.

[45] S. J. Schiff, A. Aldroubi, M. Unser, and S. Sato, "Fast wavelet transformation of eeg," *Electroencephalography and Clinical Neurophysiology*, vol. 91, no. 6, pp. 442 – 455, 1994.

[46] M. Jobert, C. Tismer, E. Poiseau, and H. Schulz, "Wavelets-a new tool in sleep biosignal analysis," *Journal of Sleep Research*, vol. 3, no. 4, p. 223–232, 1994.

[47] A. J. Izenman, *Linear Discriminant Analysis*. New York, NY: Springer New York, 2008.

[48] T. Sugi, M. Nakamura, T. Shimokawa, and F. Kawana, "Automatic detection of eeg arousals by use of normalized parameters for different subjects," *IEEE EMBS Asian-Pacific Conference on Biomedical Engineering*, 2003.

[49] R. Agarwal, "Automatic detection of micro-arousals," *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, 2005.

[50] S. Cho, J. Lee, H. Park, and K. Lee, "Detection of arousals in patients with respiratory sleep disorders using a single channel eeg," *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, 2005.

[51] O. Shmiel, T. Shmiel, Y. Dagan, and M. Teicher, "Data mining techniques for detection of sleep arousals," *Journal of Neuroscience Methods*, vol. 179, no. 2, p. 331–337, 2009.

[52] H. Mannila, H. Toivonen, and A. Inkeri Verkamo, "Discovery of frequent episodes in sequences," *KDD*, pp. 210 – 215, 1995.

[53] H. Mannila and H. Toivonen, "Discovering generalized episodes using minimal occurrences.," *KDD*, pp. 146 – 151, 1996.

[54] T. Sugi, F. Kawana, and M. Nakamura, "Automatic eeg arousal detection for sleep apnea syndrome," *Biomedical Signal Processing and Control*, vol. 4, no. 4, p. 329–337, 2009.

[55] C. K. Behera, T. K. Reddy, L. Behera, and B. Bhattacarya, "Artificial neural network based arousal detection from sleep electroencephalogram data," *2014 International Conference on Computer, Communications, and Control Technology (I4CT)*, 2014.

[56] Y. Liang, C. Leung, C. Miao, Q. Wu, and M. J. Mckeown, "Automatic sleep arousal detection based on c-elm," *2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, 2015.

[57] Wallant, Dorothée Coppieters 'T and Muto, Vincenzo and Gaggioni, Giulia and Jaspar, Mathieu and Chellappa, Sarah L. and Meyer, Christelle and Vandewalle, Gilles and Maquet, Pierre and Phillips, Christophe, "Automatic artifacts and arousals detection in whole-night sleep eeg recordings," *Journal of Neuroscience Methods*, vol. 258, p. 124–133, 2016.

[58] Fernández-Varela, Isaac and Alvarez-Estevez, Diego and Hernández-Pereira, Elena and Moret-Bonillo, Vicente, "A simple and robust method for the automatic scoring of eeg arousals in polysomnographic recordings," *Computers in Biology and Medicine*, vol. 87, p. 77–86, 2017.

[59] H. J. Caulfield and R. Haimes, "Generalized matched filtering," *Optical Pattern Recognition*, Apr 1979.

[60] G. Turin, "An introduction to matched filters," *IEEE Transactions on Information Theory*, vol. 6, no. 3, p. 311–329, 1960.

[61] M. Woertz, T. Miazhynskaia, P. Anderer, and G. Dorffner, "Automatic k-complex detection: comparison of two different approaches," *JSR*, vol. 13, p. 808, 2004.

[62] S. Salvador and P. Chan, "Toward accurate dynamic time warping in linear time and space," *Intell. Data Anal.*, vol. 11, pp. 561–580, Oct. 2007.

[63] B. Jansen, "Artificial neural nets for k-complex detection," *IEEE Engineering in Medicine and Biology Magazine*, vol. 9, no. 3, p. 50–52, 1990.

[64] V. V. Shete, S. Sonar, A. Charantimath, and S. Elgendelwar, "Detection of k-complex in sleep egg signal with matched filter fnd neural network," *International Journal of Engineering Research & Technology*, vol. 1, May 2012.

[65] S. Devuyst, T. Dutoit, P. Stenuit, and M. Kerkhofs, "Automatic k-complexes detection in sleep eeg recordings using likelihood thresholds," *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2010.

[66] A. C. da Rosa and T. Paiva, "Automatic detection of k-complexes: Validation in normals and dysthymic patients," *Sleep*, vol. 16, no. 3, p. 239, 1993.

[67] A. C. D. Rosa, B. Kemp, T. Paiva, F. H. L. da Silva, and H. A. Kamphuisen, "A model-based detector of vertex waves and k complexes in sleep electroencephalogram.," *Electroencephalography and clinical neurophysiology*, vol. 78 1, pp. 71–9, 1991.

[68] G. Zweig, "Wavelet transforms as solutions of partial differential equations," *University of California, Los Alamos, Dissertation*, Jan 1997.

[69] Z. Tang and N. Ishii, "Detection of the k-complex using a new method of recognizing waveform based on the discret wavelet transform," *IEICE Trans. Inf. Syst.*, 1995.

[70] A. Erdamar, F. Duman, and S. Yetkin, "A wavelet and teager energy operator based method for automatic detection of k-complex in sleep eeg," *Expert Systems with Applications*, vol. 39, no. 1, p. 1284–1290, 2012.

[71] R. Hamila, J. Astola, F. A. Cheikh, M. Gabbouj, and M. Renfors, "Teager energy and the ambiguity function," *IEEE Transactions on Signal Processing*, vol. 47, no. 1, p. 260–262, 1999.

[72] L. K. Krohne, R. B. Hansen, J. A. E. Christensen, H. B. D. Sorensen, and P. Jennum, "Detection of k-complexes based on the wavelet transform," *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014.

[73] T. Lajnef, S. Chaibi, J.-B. Eichenlaub, P. M. Ruby, P.-E. Aguera, M. Samet, A. Kachouri, and K. Jerbi, "Sleep spindle and k-complex detection using tunable q-factor wavelet transform and morphological component analysis," *Frontiers in Human Neuroscience*, vol. 9, 2015.

[74] T. Lajnef, C. O'reilly, E. Combrisson, S. Chaibi, J.-B. Eichenlaub, P. M. Ruby, P. Aguera, M. Samet, A. Kachouri, S. Frenette, and et al., "Meet spinky: An open-source spindle and k-complex detection toolbox validated on the open-access montreal archive of sleep studies (mass)," *Frontiers in Neuroinformatics*, Mar 2017.

[75] I. W. Selesnick and I. Bayram, "Oscillatory plus transient signal decomposition using overcomplete rational-dilation wavelet transforms," *Proc. SPIE*, vol. 7446, pp. 74460V–74460V–14, 2009.

[76] J.-L. Starck, Y. Moudden, J. Bobin, M. Elad, and D. L. Donoho, "Morphological component analysis," *Proc. SPIE*, vol. 5914, pp. 59140Q–59140Q–15, 2005.

[77] S. Yazdani, S. Fallet, and J.-M. Vesin, "A novel short-term event extraction algorithm for biomedical signals," *IEEE Transactions on Biomedical Engineering*, 2017.

[78] V. Pohl and E. Fahr, "Neuro-fuzzy recognition of k-complexes in sleep eeg signals," *Proceedings of 17th International Conference of the Engineering in Medicine and Biology Society*, p. 789–790, 1995.

[79] C. Richard and R. Lengelle, "Joint time and time-frequency optimal detection of k-complexes in sleep eeg," *Computers and Biomedical Research*, vol. 31, no. 3, p. 209–229, 1998.

[80] A. Kam, A. Cohen, A. Geva, and A. Tarasiuk, "Detection of k-complexes in sleep eeg using cd-hmm," *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2004.

[81] D. Migotina, A. Rosa, and A. Fred, "Automatic k-complex detection using hjorth parameters and fuzzy decision," *Proceedings of the 2010 ACM Symposium on Applied Computing - SAC 10*, 2010.

[82] H. Q. Vu, G. Li, N. S. Sukhorukova, G. Beliakov, S. Liu, C. Philippe, H. Amiel, and A. Ugon, "K-complex detection using a hybrid-synergic machine learning method," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, p. 1478–1490, 2012.

[83] D. A. Dean, A. L. Goldberger, R. Mueller, M. Kim, M. Rueschman, D. Mobley, S. S. Sahoo, C. P. Jayapandian, L. Cui, M. G. Morrical, and et al., "Scaling up scientific discovery in sleep medicine: The national sleep research resource," *Sleep*, vol. 39, no. 5, p. 1151–1164, 2016.

[84] E. Orwoll, J. B. Blank, E. Barrett-Connor, J. Cauley, S. Cummings, K. Ensrud, C. Lewis, P. M. Cawthon, R. Marcus, L. M. Marshall, and et al., "Design and baseline characteristics of the osteoporotic fractures in men (mros) study — a large observational study of the determinants of fracture in older men," *Contemporary Clinical Trials*, vol. 26, no. 5, p. 569–585, 2005.

[85] J. B. Blank, P. M. Cawthon, M. L. Carrion-Petersen, L. Harper, J. P. Johnson, E. Mitson, and R. R. Delay, "Overview of recruitment for the osteoporotic fractures in men study (mros)," *Contemporary Clinical Trials*, vol. 26, no. 5, p. 557–568, 2005.

[86] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of complex physiologic time series," *Physical Review Letters*, vol. 89, Jul 2002.

[87] M. Costa, A. L. Goldberger, and C. K. Peng, "Multiscale entropy analysis of biological signals," *PHYSICAL REVIEW*, Feb 2005.

[88] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology - Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–H2049, 2000.

[89] S. Mariani, A. F. T. Borges, T. Henriques, A. L. Goldberger, and M. D. Costa, "Use of multiscale entropy to facilitate artifact detection in electroencephalographic signals," *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015.

[90] "Sleep data - national sleep research resource - nsrr."

[91] J. Buckelmüller, H.-P. Landolt, H. Stassen, and P. Achermann, "Trait-like individual differences in the human sleep electroencephalogram," *Neuroscience*, vol. 138, no. 1, pp. 351 – 356, 2006.

[92] D. Brunner, R. Vasko, C. Detka, J. Monahan, C. R. Iii, and D. Kupfer, "Muscle artifacts in the sleep eeg: Automated detection and effect on all-night eeg power spectra," *Journal of Sleep Research*, vol. 5, no. 3, p. 155–164, 1996.

[93] J. Pan and W. J. Tompkins, "A real-time qrs detection algorithm," *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, p. 230–236, 1985.

[94] F. Gustafsson, "Determining the initial states in forward-backward filtering," *IEEE Transactions on Signal Processing*, vol. 44, pp. 988–992, Apr 1996.

[95] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-time signal processing*. Prentice-Hall, 2014.

[96] E. A. Ashley and J. Niebauer, *Cardiology Explained*. Remedica Medical Education and Publishing, 2004.

[97] S. Mariani, L. Tarokh, I. Djonlagic, B. Cade, M. Morrical, K. Yaffe, K. Stone, K. Loparo, S. Purcell, and D. Aeschbach, "0763 automated pipeline for spec-

tral analysis of eeg data: The national sleep research resource tool," *Sleep*, vol. 40, no. 1, 2017.

[98] M. Salinsky, T. Sutula, and D. Roscoe, "Representation of sleep stages by color density spectral array," *Electroencephalogry and Clinical Neurophysiology*, vol. 66, p. 579–582, 1987.

[99] M. Salinsky, S. Goins, T. Sutula, D. Roscoe, and S. Weber, "Comparison of sleep staging by polygraph and color density spectral array," *Sleep*, p. 131–138, Jan 1988.

[100] R. N. Bracewell, *The Fourier transform and its applications*. McGraw Hill, 3 ed., 2000.

[101] J. Allen, "Short term spectral analysis, synthesis, and modification by discrete fourier transform," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 25, p. 235–238, 1977.

[102] J. F. Kaiser, "Nonrecursive digital filter design using the i-sinh window function," *Proceedings of the 1974 IEEE International Symposium on Circuits and Systems*, p. 20–23, Apr 1974.

[103] D. Thomson, "Spectrum estimation and harmonic analysis," *Proceedings of the IEEE*, vol. 70, no. 9, p. 1055–1096, 1982.

[104] T. Bronez, "On the performance advantage of multitaper spectral analysis," *IEEE Transactions on Signal Processing*, vol. 40, no. 12, p. 2941–2946, 1992.

[105] D. B. Percival and A. T. Walden, "Spectral analysis for physical applications: Multitaper and conventional univariate techniques," *Cambridge University Press*, p. 456–545, 1993.

[106] D. Slepian, "Prolate spheroidal wave functions, fourier analysis, and uncertainty-v: The discrete case," *Bell System Technical Journal*, vol. 57, p. 1371–1430, Jun 1978.

[107] M. J. Prerau, R. E. Brown, M. T. Bianchi, J. M. Ellenbogen, and P. L. Purdon, "Sleep neurophysiological dynamics through the lens of multitaper spectral analysis," *Physiology*, vol. 32, p. 60–92, Dec 2016.

[108] T. Fawcett, "An introduction to roc analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861 – 874, 2006.

[109] A. P. Bradley, "The use of the area under the roc curve in the evaluation of machine learning algorithms," *Pattern Recognition*, vol. 30, no. 7, pp. 1145 – 1159, 1997.

Appendices

Chapter A

Invalid Subjects

From the all the data, three subjects from each dataset were considered invalid and excluded in the subset II. In this section, it will be reported the main reasons for the exclusion of this subjects.

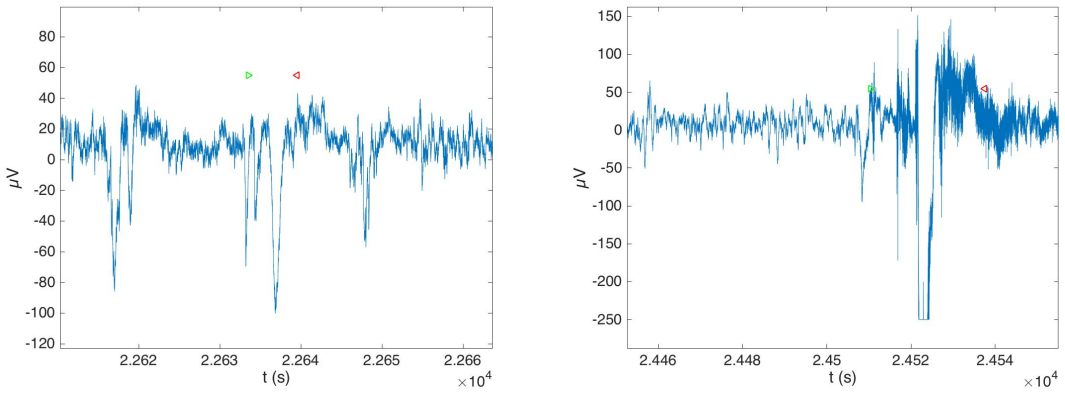
A.1 MrOS

The three subjects that were excluded from subset II from the MrOS dataset were subjects, 5, 8, and 18. The validity of these subjects will be inspected in the next sections.

A.1.1 Subject 5

The inspection of the signal, showed no significant noise, neither was the power spectrum of the signal in any way irregular. So it was inspected the visual classification performed by the expert. The labels seemed very unusual, most of the

arousal classification were dislocated randomly from the actual raise of frequency of the EEG signal. An example is shown in Figure A.1.



(a) **(b)**

Figure A.1: Two examples of the visual scoring of subject 5 from MrOS dataset. The arousal is marked with their start (green arrow), and their end (red arrow).

In Figure A.1a as it can be seen the label of the arousal is not an arousal but actually a K-complex. In Figure A.1b, what was visually scored as an arousal was actually an artifact. The artifact detectors could solve the problem of the second example, but not the first. The visual scoring of this subject was not valid so although the classifier could detect this transient event, the performance evaluation would not be satisfying.

A.1.2 Subject 8

Subject 8 had a lot of noise. The power spectrum was inspected and it can be seen in Figure A.2.

In the Power spectrum, it is possible to see multiple peaks at different frequencies. The peak at 60hz is due to the powerline interference. This is due to differences in the electrode impedances and the stray currents through the patient. Because the signal was not filtered before the analogic-to-digital conversion, it is possible to see the harmonic frequencies mirrored in the power spectrum. This corrupts the signal and these peaks in the time domain represent a constant periodic interference at these frequencies.

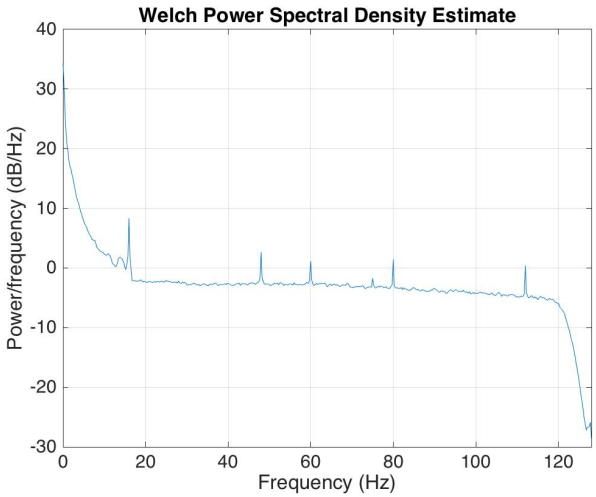
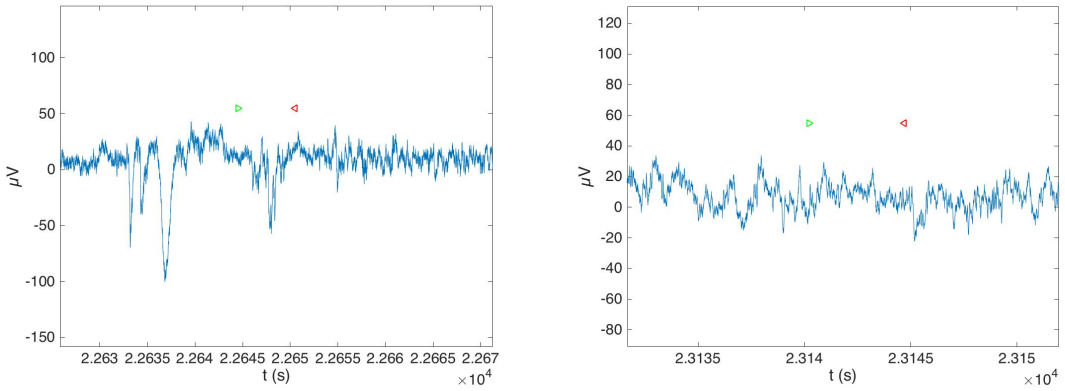


Figure A.2: The power spectrum estimate of subject 8 calculated using Welch method.

A.1.3 Subject 18

The inspection of the signal, showed no significant noise, neither was the power spectrum of the signal in any way irregular. So it was inspected the visual classification performed by the expert like before. From all the visual scores half of them were classified as an artifact, turning out not to be valid. An example of the visual scores is shown in Figure A.3.



(a) **(b)**

Figure A.3: Two examples of the visual scoring of subject 18 from MrOS dataset. The arousal is marked with their start (green arrow), and their end (red arrow).

In Figure A.3a the visual score is associated to what could be an artifact. In Figure A.3b second example of a the visual score, it is not noticeable any increase of frequency, this scoring may be done to other EEG channel.

The visual scoring of this subject was not valid so although the classifier could detect this transient event, the performance evaluation would not be satisfying.

A.2 MESA

The three subjects that were excluded from subset *II* from the MrOS dataset were subjects, 5, 6, and 16. The validity of these subjects will be inspected in the next sections.

A.2.1 Subject 5

The power spectrum of subject 5 was inspected and it can be seen in the Figure A.4.

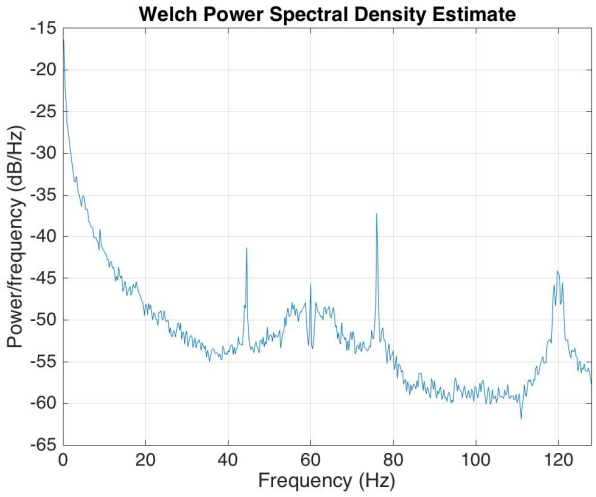


Figure A.4: The power spectrum estimate of subject 5 calculated using Welch method.

From the analyse of the power spectrum it is possible to conclude that the signal is full of noise. There is a peak at approximately 43 Hz, for unknown reasons. It is very likely that the signal was badly acquired generating this interferences and invalidating the signal.

A.2.2 Subject 6

This subject had a very noisy signal and when the artifact detector was applied to it, the amount of detections was of an higher order of magnitude from the rest of the

subjects, suggesting high noise levels. The power spectrum was inspected anyway and is reported in Figure A.5.

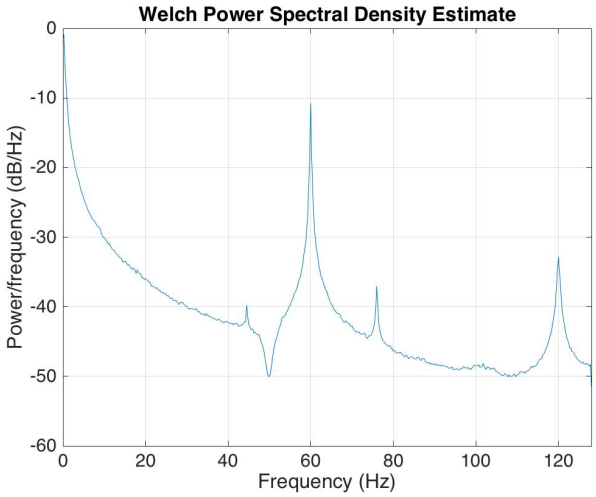


Figure A.5: The power spectrum estimate of subject 6 calculated using Welch method.

The spectrum has an odd shape, the power interference is very strong but the signal is completely corrupted by noise. For this reasons it was excluded from subset *II*.

A.2.3 Subject 18

The power spectrum of subject 18 is reported in Figure A.6.

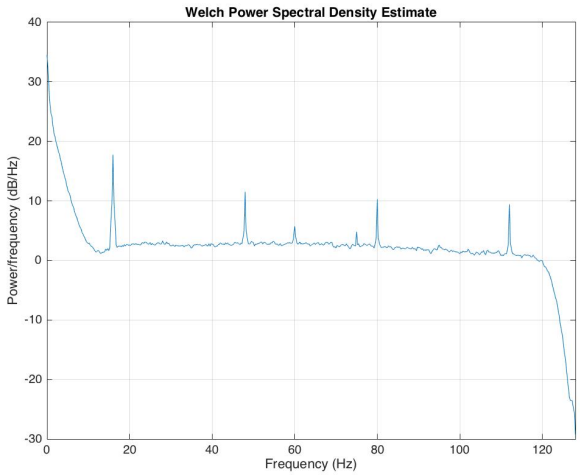


Figure A.6: The power spectrum estimate of subject 18 calculated using Welch method.

In the Power spectrum, it is possible to see multiple peaks at different frequencies. The peak at 60hz is due to the powerline interference. Besides the signal was not filtered before the analogic-to-digital conversion and it is possible to see the harmonic frequencies mirrored in the power spectrum. This corrupts the signal so it was excluded from the dataset.

Chapter B

ROCs

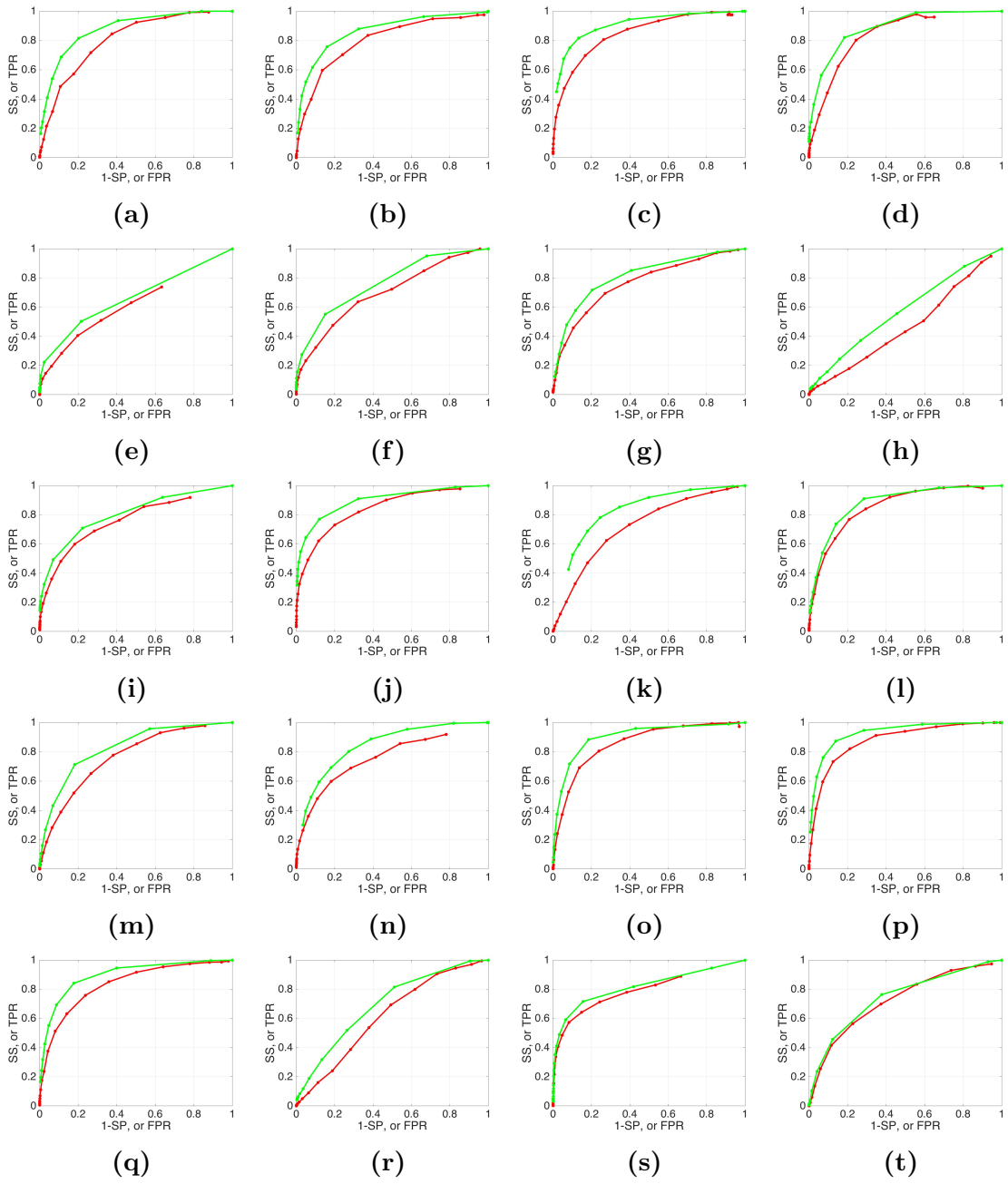


Figure B.1: The roc curves for each subject of the MrOS dataset for the two detectors Multitaper (green) and Spectrogram (red).

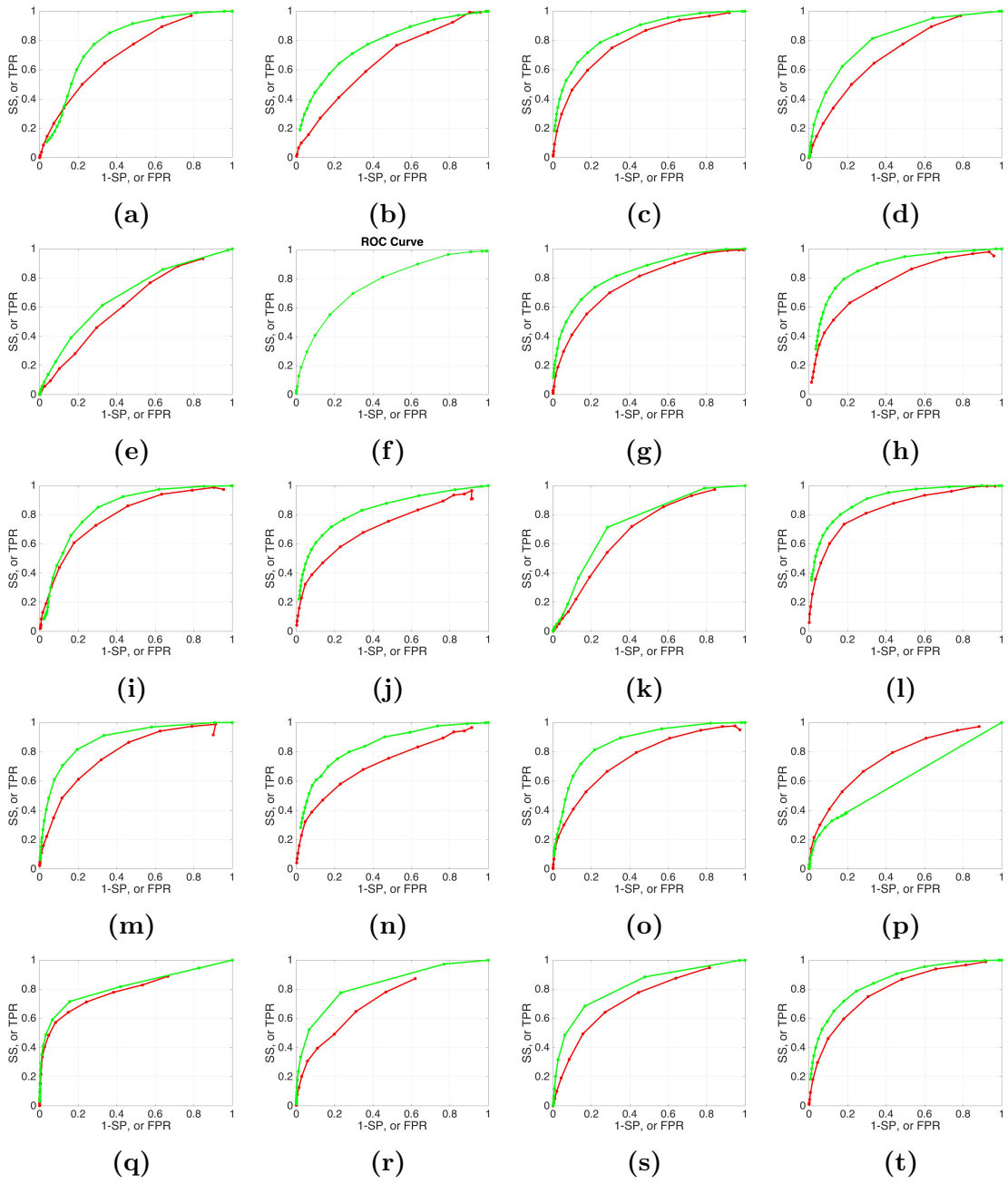


Figure B.2: The roc curves for each subject of the MESA dataset for the two detectors Multitaper (green) and Spectrogram (red).

Chapter C

Statistical Results

Table C.1: Number of arousals for all subjects from both datasets, before and after the validation.

	MrOS		*MESA*	
Subject	*Visual Classification*	*Valid*	*Visual Classification*	*Valid*
1	54	33	197	114
2	74	47	169	145
3	105	53	189	105
4	87	34	97	74
5	175	89	161	109
6	106	68	89	46
7	83	54	124	92
8	99	54	264	223
9	133	65	124	63
10	182	111	194	154
11	170	95	334	236
12	173	120	89	57
13	238	177	122	83
14	74	62	43	17
15	186	67	124	64
16	193	148	145	106
17	164	109	88	67
18	273	205	115	34
19	253	155	179	151
20	85	64	116	49

Table C.2: The statistical results of the subjects from MrOS dataset using spectrogram as an arousal detector.

Subject	*AUC*	*SP*	*SS*	*ACC*	*Ideal Threshold*
1	0.7868	0.6930	0.7214	0.6947	7
2	0.7985	0.8415	0.7133	0.8384	9
3	0.8545	0.7010	0.8380	0.7041	9
4	0.7342	0.7333	0.8139	0.7342	5
5	0.5741	0.6472	0.4989	0.6408	2
6	0.7072	0.6379	0.6993	0.6400	5
7	0.8031	0.6946	0.7824	0.6970	7
8	0.4601	0.4662	0.4481	0.4658	6
9	0.7590	0.7821	0.6780	0.7788	6
10	0.8440	0.7731	0.7779	0.7733	7
11	0.7193	0.6855	0.6581	0.6844	10
12	0.8538	0.7673	0.7946	0.7684	6
13	0.7644	0.7001	0.6944	0.6996	5
14	0.7514	0.6044	0.7807	0.6163	9
15	0.8596	0.7324	0.8368	0.7482	7
16	0.8898	0.8572	0.7844	0.8511	9
17	0.8568	0.7222	0.8046	0.7364	7
18	0.5950	0.4698	0.7117	0.4950	5
19	0.8505	0.7593	0.7973	0.7618	5
20	0.7262	0.7379	0.5861	0.7331	7

Table C.3: The statistical results of the subjects from MESA dataset using spectrogram as an arousal detector.

Subject	*AUC*	*SP*	*SS*	*ACC*	*Ideal Threshold*
1	0.7069	0.7019	0.6167	0.6947	-49
2	0.7752	0.7216	0.7696	0.8384	-47
3	0.7897	0.6931	0.7498	0.7041	-48
4	0.7131	0.6615	0.6458	0.7342	-51
5	0.6235	0.5648	0.6235	0.6408	-51
6	0.4226	0.5870	0.3601	0.6400	-56
7	0.7653	0.7051	0.6993	0.6970	-50
8	0.7696	0.7880	0.6300	0.4658	-48
9	0.7862	0.7080	0.7278	0.7788	-51
10	0.7526	0.6508	0.6785	0.7733	-48
11	0.6875	0.5896	0.7198	0.6844	-50
12	0.8379	0.8173	0.7363	0.7684	-47
13	0.7874	0.6798	0.7461	0.6996	-50
14	0.7565	0.6407	0.6758	0.6163	-50
15	0.7859	0.7173	0.6659	0.7482	-56
16	0.5122	0.8078	0.3824	0.8511	-53
17	0.7958	0.7568	0.7132	0.7364	-55
18	0.7264	0.6883	0.6486	0.4950	-54
19	0.7347	0.7285	0.6430	0.7618	-53
20	0.7733	0.6676	0.7481	0.7331	-48

Table C.4: The statistical results of the subjects from MrOS dataset using Multitaper as an arousal detector.

Subject	*AUC*	*SP*	*SS*	*ACC*	*Ideal Threshold*
1	0.8767	0.8196	0.8242	0.8191	32
2	0.8637	0.8160	0.7763	0.8269	38
3	0.9027	0.8530	0.8327	0.8581	50
4	0.8864	0.8158	0.8202	0.8159	20
5	0.6776	0.7832	0.5625	0.7097	10
6	0.7657	0.8489	0.5510	0.7107	20
7	0.8128	0.7960	0.7176	0.7929	30
8	0.5758	0.5678	0.5347	0.5669	21
9	0.8016	0.7193	0.7771	0.7219	18
10	0.8954	0.8548	0.7979	0.8515	28
11	0.8088	0.7541	0.7816	0.7555	60
12	0.8799	0.7873	0.8457	0.7912	24
13	0.8287	0.7740	0.7754	0.7742	18
14	0.8359	0.7262	0.8021	0.7320	50
15	0.9029	0.8395	0.8643	0.8422	32
16	0.9277	0.8604	0.8732	0.8617	40
17	0.8955	0.8365	0.8330	0.8362	31
18	0.6944	0.6054	0.6978	0.6163	24
19	0.8766	0.8225	0.8161	0.8220	21
20	0.7427	0.6642	0.7336	0.6991	21

Table C.5: The statistical results of the subjects from MESA dataset using spectrogram as an arousal detector.

Subject	*AUC*	*SP*	*SS*	*ACC*	*Ideal Threshold*
1	0.7854	0.7290	0.7737	0.7310	72
2	0.7752	0.7216	0.7696	0.7200	102
3	0.8503	0.7687	0.7764	0.7692	92
4	0.8082	0.7629	0.7183	0.7586	35
5	0.6824	0.6741	0.6394	0.6448	29
6	0.4026	0.5970	0.3401	0.5080	10
7	0.8326	0.7772	0.7462	0.7715	69
8	0.8637	0.8127	0.7980	0.8116	89
9	0.8346	0.7771	0.8244	0.7303	60
10	0.8306	0.7940	0.7399	0.7904	76
11	0.7422	0.6990	0.7543	0.7003	19
12	0.9002	0.8459	0.7976	0.8435	92
13	0.8818	0.7936	0.8293	0.7959	49
14	0.8376	0.7641	0.7809	0.7645	86
15	0.8639	0.8012	0.7999	0.8012	62
16	0.6113	0.8062	0.3850	0.6180	10
17	0.8184	0.8370	0.7266	0.8244	29
18	0.8286	0.7354	0.8035	0.7382	19
19	0.8179	0.7458	0.7776	0.7486	26
20	0.8849	0.8479	0.7995	0.8465	35

Table C.6: The number of K-complex detected at each sleep stage using matched filters for each subject of the MrOS dataset.

	# of K-complex detections				
Subject	*nREM1*	*nREM2*	*nREM3*	*nREM4*	*REM*
1	136	1043	399	0	574
2	238	3412	313	0	684
3	51	229	17	0	66
4	31	255	19	0	60
5	31	91	8	0	93
6	103	3544	206	11	395
7	19	168	8	0	132
8	120	2130	21	0	641
9	34	160	17	3	137
10	232	1822	174	0	933
11	320	2335	1050	0	435
12	48	244	15	0	84
13	2	29	0	0	17
14	7	71	0	0	0
15	159	783	5	0	18
16	39	439	50	0	26
17	88	1078	21	0	516
18	4	121	24	0	39
19	261	319	1	0	273
20	12	85	5	0	6

Table C.7: The number of K-complex detected at each sleep stage using matched filters for each subject of the MESA dataset.

	# of K-complex detections				
Subject	*nREM1*	*nREM2*	*nREM3*	*nREM4*	*REM*
1	986	4470	148	0	371
2	456	3110	648	0	1137
3	36	78	31	0	30
4	145	1127	695	0	0
5	359	2063	137	0	200
6	94	4077	2789	0	481
7	537	5160	477	0	753
8	35	162	64	0	0
9	640	5190	227	0	886
10	220	1315	648	0	405
11	386	1253	0	0	200
12	404	1979	405	0	527
13	171	1246	810	0	55
14	114	1477	868	0	339
15	770	3122	2075	0	916
16	112	754	254	0	146
17	85	2086	770	0	492
18	148	1871	39	0	260
19	504	3984	726	0	367
20	389	3456	700	0	638